

Table 1

BCA4 DNA sequence (SEQ ID NO:1)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset Accession #: D13666; Nucleic Acid Accession #: NM_006475; Coding sequence: 12-2522 (start and stop codons underlined)

AGAGACTCAA	GATGATTCCC	TTTTTACCCA	TGTTTCTCT	ACTATTGCTG	CTTATTGTTA	60
ACCCTATAAA	CGCCAACAAT	CATTATGACA	AGATCTTGGC	TCATAGTCGT	ATCAGGGGTC	120
GGGACCAAGG	CCCAATGTC	TGTGCCCTTC	AACAGATTTT	GGGCACCAAA	AAGAAATACT	180
TCAGCACTTG	TAAGAACCTGG	TATAAAAAGT	CCATCTGTGG	ACAGAAAAGC	ACTGTTTAT	240
ATGAATGTTG	CCCTGGTTAT	ATGAGATATGG	AAGGAATGAA	AGGCTGCCCA	GCAGTTTTC	300
CCATTGACCA	TGTTTATGGC	ACTCTGGCA	TCGTGGGAGC	CACCAACAGC	CAGCGCTATT	360
CTGACGCCCTC	AAAATGAGG	GAGGAGATCG	AGGGAAAGGG	ATCCTTCACT	TACTTTGAC	420
CGAGTAATGA	GGCTTGGGAC	AACCTGGATT	CTGATATCCG	TAGAGGTTTG	GAGAGCAACG	480
TGAATGTTG	ATTACTGAAT	GCTTTACATA	GTACATGAT	TAATAAGAGA	ATGTTGACCA	540
AGGACTTAA	AAATGGCATG	ATTATTCTT	CAATGTATAA	CAATTGGGG	CTTTTCATTA	600
ACCATTATCC	TAATGGGGTT	GTCACTTTA	ATTGTGCTCG	AATCATCCAT	GGGAACAGA	660
TTGCAACAAA	TGGTGTGTC	CATGTCATTG	ACCGTGTGCT	TACACAAATT	GGTACCTCAA	720
TTCAAGACTT	CATTGAAGCA	GAAGATGACC	TTTCATCTTT	TAGAGCAGCT	GCCATCACAT	780
CGGACATATT	GGAGGCCCTT	GGAGAGAGCC	GTCACTTCAC	ACTCTTGTGCT	CCACCAATG	840
AGGCTTTGAA	GAAACATTC	CGAGGTGTT	TAGAAAGGTT	CATGGAGAC	AAAGTGCCTT	900
CCGAAGCTCT	TATGAGTAC	CACATCTAA	ATACTCTCA	GTGTTCTGAG	TCTATTATGG	960
GAGGAGCAGT	CTTTGAGACG	CTGGAAGGAA	ATACAATTGA	GATAGGATGT	GACGGTGACA	1020
GTATAACAGT	AAATGGAATC	AAAATGGTGA	ACAAAAAGGA	TATTGTGACA	AATAATGGT	1080
TGATCCATT	GATTGATCAG	GTCCTAATTG	CTGATTCTGC	CAAACAAGGT	ATTGAGCTGG	1140
CTGGAAACAA	GCAAAACACC	TTCACGGATC	TTGCGGCCA	ATTAGGCTTG	GCATCTGCTC	1200
TGAGGCCAGA	TGGAGAATAC	ACTTTGCTGG	CACCTGTGAA	TAATGCATT	TCTGATGATA	1260
CTCTCAGCAT	GGTTTCAGCGC	CTCCCTTAAT	TAATTCTGCA	GAATCACATA	TTGAAAGTAA	1320
AAGTTGGCCT	TAATGAGCTT	TACAACGGGC	AAATACTGGA	AACCATCGGA	GGCAAACAGC	1380
TCAGAGTCTT	CGTATATCGT	ACAGCTGTCT	GCATTGAAAA	TTCATGCTG	GAGAAAGGGA	1440
CTAAGCAAGG	GAGAACACGGT	GCGATTCACA	TATTCGGCGA	GATCATCAAG	CCAGCAGAGA	1500
ATACCTCCA	TGAAAGATTA	AAACAAGATA	AGCGCTTTAG	CACCTTCCTC	AGCCTACTTG	1560
AAAGCTGCAGA	CTTGAAGAG	CTCCTGACAC	AAACCTGGAGA	CTGGACATTA	TTTGTGCCAA	1620
CCAATGATGC	TTTTAAGGGG	ATGACTAGTG	AAGAAAAAGA	AATTCTGATA	CGGGACAAAA	1680
ATGCTCTTCA	AAACATCATT	CTTTATCACC	TGACACCAGG	AGTTTTCAT	GGAAAAGGAT	1740
TTGAACCTGG	TGTTACTAAC	ATTTTAAAGA	CCACACAAGG	AAGCAAATTC	TTTCTGAAAG	1800
AAGTAATATGA	TACACTCTG	GTGAATGAAT	TGAAATCAA	AGAATCTGAC	ATCATGACAA	1860
CAAATGGTGT	AATTCTGTT	GTAGATAAAC	TCCTCTATCC	AGCAGACACA	CCTGTTGGAA	1920
ATGATCAACT	GCTGGAAATA	CTTAATAAAAT	TAATCAAATA	CATCCAATT	AAGTTTGTT	1980
GTGGTAGCAC	CTTCAAAGAA	ATCCCCGTGA	CTGCTATAC	AACTAAAATT	ATAACCAAAG	2040
TTGTGGAACC	AAAAATTAAA	GTGATTGAAG	GCAGCTTCA	GCCTATTATC	AAAATGAAAG	2100
GACCCACACT	AACAAAAGTC	AAAATTAAGA	GTGAACCTGA	ATTCAAGACTG	ATTAAAAGAAG	2160
GTGAAACAAAT	AACTGAAGTG	ATCCATGGAG	AGCCAATTAT	AAAAAAATAC	ACCAAAATCA	2220
TTGATGGAGT	GCCTGTGAA	ATAACTGAAA	AAGAGACACG	AGAAGAACGA	ATCATTACAG	2280
GTCCTGAAAT	AAAATACACT	AGGATTCTA	CTGGAGGTGG	AGAAACAGAA	GAAACTCTGA	2340
AGAAATTGTT	ACAAGAAGAG	GTCACCAAGG	TCACCAAATT	CATTGAAGGT	GGTGATGGTC	2400
ATTTTATTGTA	AGATGAAGAA	ATTAAGAACG	TGCTTCAGGG	AGACACACCC	GTGAGGAAGT	2460
TGCAAGCCAA	AAAAAAAGTT	CAAGGTCTA	GAAGACGATT	AAGGGAAAGGT	CGTTCTCAGT	2520
<u>GAAAATCCAA</u>	<u>AAACCAAGAA</u>	<u>AAAATGTTA</u>	<u>TACAACCCCTA</u>	<u>AGTCATAAAC</u>	<u>CTGACCTTAG</u>	<u>2580</u>
AAAATTGTGA	GAGCCAAGTT	GACTTCAGGA	ACTGAAACAT	CAGCACAAG	AAGCAATCAT	2640
CAAATAATT	TGAACACAAA	TTTATATT	TTTTTCTGA	ATGAGAAAC	TGAGGGAAAT	2700
TGTGGAGTTA	GCCTCTGTC	GTAAAGGAAT	TGAAGAAAAT	ATAACACCTT	ACACCCCTTT	2760
TCATCTGAC	ATTAAGGAT	TCGGCTAAT	TTGGAATCTA	TTAGGAAAAA	ATCCTTGTC	2820
CCAGATTCT	TACAATTCAA	ATCGAAGAGT	TGTGAACCTG	TATCCCATTG	AAAAGACCGA	2880
GCCTTGATG	TATGTTATGG	ATACATAAAA	TGCAACGAAG	CCATTATCTC	TCCATGGGAA	2940
GCTAAGTTAT	AAAAATAGGT	GCTTGGTGT	CAAAACTTTT	TATATCAAAA	GGCTTTGCAC	3000
ATTTCATAT	GAGTGGTTT	ACTGGTAAAT	TATGTTATT	TTTACAACTA	ATTTTGACT	3060
CTCAGAATGT	TTGTCATATG	CTTCTGCAA	TGCAATTTT	TTAATCTCAA	ACGTTCAAT	3120
AAAACATT	TTCAGATATA	AAAGAGATTA	CTTCAAATTG	AGTAATTCA	AAAACCTAA	3180
GATTAAAGTT	AAAAAGTGGT	TTGGACTTGG	GAA			

BCA4 Protein sequence (SEQ ID NO:2)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset Accession #: D13666; Protein Accession #: NP_006466; Predicted Signal sequence: 1-21; TM domains: none; PFAM domains: fasciclin_domains: 94-232, 234-367, 496-630; Summary: a secreted protein involved in adhesion and osteoblast development; may participate in preferential metastasis of breast cancer to bone.

MIPFLPMFSL	LLLLIVNPIN	ANNHYDKILA	HSRIRGRDQG	PNVCALQQIL	GTKKKYFSTC	60
KNWYKKSICG	OKTTVLYECC	PGYMRMEGMK	GCPAVLPIDH	VYGTGLIVGA	TTTQRYSDas	120
KLREEIEKGK	SFTYFAPSNE	AWDNLDSDIR	RGLESNVNVE	LLNALHSHMI	NKRMLTKDLK	180
NGMIIIPSMYN	NLGLFINHYP	NGVVTVN CAR	IIHGNQIATN	GVVVHVIDRVL	TQIGTSIQDF	240
IEAEEDDLSSF	RAAAITSIDL	EALGRDHFT	LFAPTNEAFE	KLPRGVLERF	MGDKVASEAL	300
MKYHILNTLQ	CSESIMGGAV	FETLEGNTI	IGCDGDSITV	NGIKMVNKKD	IVTNNGVIHL	360
IDQVLIIPDSA	KQVIELAGKQ	QTTFTDLVAQ	LGLASALRPD	GEYTLLAPVN	NAFSDDTLSM	420

VQRLKLILQ NHILKVKVGL NELYNGQILE TIGGKQLRWF VYRTAVCIEN SCMEKGSKQG	480
RNGAIHIFRE IIKPAEKSILH EKLKQDKRFS TFLSLLEAAD LKELLTQPGD WTLFVPTNDA	540
FKGMTSEEKE ILIRDKNALQ NIILYHLTPG VFIGKGFEPEG VTNLKTTQG SKIFLKEVND	600
TLLVNELKSK ESDIMTTNGV IHVVDKLLYP ADTPVGNDQL LEILNKLIKY IQIKFVRGST	660
FKEIPVTVYT TKIITKVVEP KIKVIEGSLQ PIKTEGPTL TKVKIEGEPE FRLIKEGETI	720
TEVIHGEPII KYTKIIDGV PVEITEKETR EERIITGPEI KYTRISTGGG ETEETLKKLL	780
QEevtKVTKF IEQGDGHLF DEEIKRLLQG DTPVRKLQAN KKVGQSRRRL REGRSQ	

BCA7 DNA sequence (SEQ ID NO:3)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset Accession #: Z29083; Nucleic Acid Accession #: NM_006670; Coding sequence: 85-1347 (start and stop codons underlined)

CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCCCAGCCC	60
AGCTCCGGGG AAACCGGAGC CGCGATGCCG GGGGGGTGCT CCCGGGGGCC CGCCGCCGGG	120
GACGGGGCGT TGCGCTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCCCTCGTCT	180
TCTCCCACCT CCTCGGCATC CTCCCTTC CTCCTCGGCG CGTTCCTGGC TTCCGCGGTG	240
TCCGCCAGC CCCCCCTGCC GGACCACTG CCGCGCTGT GCGAGTGCTC CGAGGCAGCG	300
CCGACAGTCA AGTGCCTAA CCGCAATCTG ACCGAGGTGC CCACGGACCT GCCCGCCTAC	360
GTGCGCAACC TCTTCCCTAC CGGCAACCCG CTGGCCGTGC TCCCTGCCGG CGCCTTCGCC	420
CGCCGGCGC CGCTGGCGGA GCTGGCGCG CTCAACCTCA CGGGCAGCGG CCTGGACGAG	480
GTGCGCCGGG GCGCTTCGGA GCATCTGCCG AGCTCGGCC AGCTGACCT CAGCCACAAC	540
CCACTGGCCG ACCTCAGTC CCTCGCTTC TCGGGCAGCA ATGCCAGCGT CTCGGCCCCC	600
AGTCCCCTTG TGGAACTGAT CCTGAACAC ATCGTCCCCC CTGAAGATGA GCGGCAGAAC	660
CGGAGCTTCG AGGGCATGGT GGTGGCGGCC CTGCTGGCGG GCCGTGCACT GCAGGGGCTC	720
CGCCGCTTGG AGCTGGCACG CAACCACTTC CTTAACCTGC CGCGGGATGT GCTGGCCCAA	780
CTGCGCCAGCC TCAGGACCTC GGACTTAAGT AATAATTGCG TGTTGAGCCT GACCTACGTG	840
TCCTTCGCA ACCTGACACA TCTAGAAAGC CTTCACCTGG AGGACAATGC CCTCAAGGTC	900
CTTCACAATG GCACCCCTGC TGAGTTGCAA GGTCTACCCC ACATTAGGGT TTTCTGGAC	960
AACAATCCCT GGGTCTGCGA CTGCCACATG GCACAGATGG TGACCTGGCT CAAGGAAACA	1020
GAGGTAATGTC AGGGCAAAGA CCGGCTCACC TGTGATATC CGGAAAAAAT GAGGAATCGG	1080
GTCTCTTGGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTC CCCATCCCTG	1140
CAAACCTCTT ATGTCTTCT GGGTATTGTT TTAGGCCCTGA TAGGCGCTAT TTTCTCTTG	1200
GTTTTGATT TGAACCGAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCTG	1260
AGGGATCACA TGGAAAGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA	1320
AACCTCAGTT CTAACTCGGA TGTCTGAGAA ATATTAGAGG ACAGACCAAG GACAACCTCG	1380
CATGAGATGT AGACTTAAGC TTATACCTCA CTAGGCTTG TCCACTTTCA TCTCCACTA	1440
TAGATACAAAC GGACTTGCAC TAAAGCAGT GAAGGGGATT TGCTTCTTG TTATGTAAG	1500
TTTCTCGGTG TGTTCTGTTA ATGTAAGACG ATGAACAGTT GTGTATAGTG TTTTACCTC	1560
TTCTTTTCT TGGAACCTCT CAACACGTAT GGAGGGATT TTCAGGTTTC AGCATGAACA	1620
TGGGCTTCTT GCTGCTGTC TCTCTCTCG TACAGTTCAA GGTGTAGCAA GTGTACCCAC	1680
ACAGATAGCA TTCAACAAAA GCTGCTCAA CTTTTCGAG AAAAATACTT TATTCAAAA	1740
TATCAGTTT ATTCTCATGT ACCTAAGTT TGAGAAAAAT AATTGCATCC TATAAACTGC	1800
CTGCAGACGT TAGCAGGCTT TCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA	1860
AGAGCATGCT TACATTTAC TGTTCTGCAT ATTACAAAAA ATAACATTGCA ACTTCATAAC	1920
TTCTTGACA AAGTAAATTA CTTTTTGAT TGCAAGTTAT ATGAAAATGT ACTGATTTT	1980
TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAGAAAAA AAAAATAAAG	2040
ATTCTTAAAAA GAA	

BCA7 Protein sequence (SEQ ID NO:4)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset Accession #: Z29083; Protein Accession #: NP_006661; Predicted Signal sequence: 1-32; Predicted TM domains: 357-373; PFAM domains: leucine rich repeats: 61-90, 119-142, 143-166, 235-258, 259-282, 294-345; Summary: a type 1a TM protein of unknown function, detected in multiple cancers, with highest expression in breast cancer.

MPGGCSRGPAGDGRLRLAR LALVLLGWVS SSSPTSSASS FSSSAPPLAS AVSAQPPPLD	60
QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLAVLPAGA FARRPPLAEL	120
AALNLSGSRL DEVRAFAEH LPSLRQLDLS HNPLADLSPF AFSGSNASVS APSPLVELIL	180
NHIVPPEDER QNRSFEGMVV AALLLAGRALQ GLRRLELASN HFLYLPRDVL AQLPSLRHLD	240
LSNNSLVSLT YVSFRNLTHL ESLHLEDNAL KVLHNGLTAE LQGLPHIRVF LDNNNPWVCDC	300
HMDADMVTWLK ETEVVQGKDR LTCAYPEKMR NRVLLELNSA DLDCDPILPP SLQTSYVFLG	360
IVLALIGAIF LLVLYLNRKG IKKWMHNIRD ACRDHMEGYH YRYEINADPR LTNLSSNSDV	

BCX5 DNA sequence (SEQ ID NO:5)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Nucleic Acid Accession #: AF160477; Coding sequence: 225-1757 (start and stop codons underlined)

GGGGAGCTCG GAGCTCCCGA TCACGGCTTC TTGGGGTAG CTACGGCTGG GTGTGTAGAA	60
CGGGGCCGGG GCTGGGGCTG GGTCCCCTAG TGAGACCCAA GTGCGAGAGG CAAGAACTCT	120
GCAGCTTCTCT GCCTTCTGGG TCAGTCTCTT ATTCAAGTCT GCAGCCGCT CCCAGGGAGA	180
TCTCGGTGGA ACTTCAGAAA CGCTGGGAGC TCTGCCTTTC AACCATGCCCG CTGTCCCTGG	240
GAGCCGAGAT GTGGGGGCTT GAGGGCTGGC TGCTGCTGCT GCTACTGCTG GCATCATTAA	300
CAGGCCGGTG CCCCGGGGT GAGCTGGAGA CCTCAGACCT GGTAACTGTG GTGCTGGGCC	360
AGGACGAAA ACTGCCCTGC TTCTACCGAG GGGACTCCGG CGAGCAAGTG GGGCAAGTGG	420

CATGGGCTCG	GGTGGACGCG	GGCGAAGGCG	CCCAGGAAC	AGCGCTACTG	CACTCCAAAT	480
ACGGGCTTC	TGTGAGCCG	GCTTACGAGG	GCCGCGTGG	GCAGCCGCCG	CCCCCACGCA	540
ACCCCTGGA	CGGCTCAGTG	CTCCCTCGCA	ACGCACTGCA	GGCGGATGAG	GGCGAGTACG	600
AGTGCCTGG	CAGCACCTTC	CCCGCCCGG	GCTTCCAGGC	GCGGCTGCG	CTCCGAGTGA	660
TGGTGCCTCC	CCTGCCCCA	CTGAATCTG	GTCCAGCACT	AGAACAGGGC	CAGGGCCTGA	720
CCCTGGCAGC	CTCCTGCACA	GCTGAGGGCA	GCCCCAGCCCC	CAGCGTGACC	TGGGACACGG	780
AGGTCAAAGG	CACAACGTCC	AGCCGTTCT	TCAAGCACTC	CCGCTCTGCT	GGCGTCACCT	840
CAGAGTTCCA	CTTGGTGCCT	AGCCGAGCA	TGAATGGGCA	GCCACTGACT	TGTGTGGTGT	900
CCCATCCCTGG	CCTGCTCCAG	GACCAAAAGG	TCACCCACAT	CCTCCACGCT	TCCTTCCTTG	960
CTGAGGCTC	TGTGAGGGC	CTTGAAAGAC	AAAATCTGTG	GCACATTGGC	AGAGAACGGG	1020
CTATGCTCAA	GTGCGCTGAGT	GAAGGGCAGC	CCCCCTCCCTC	ATACAACCTGG	ACACGGCTGG	1080
ATGGGCCCTCT	CCCCAGTGGG	GTACGACTGG	ATGGGGACAC	TTTGGGCTTT	CCCCCACTGA	1140
CCACTGAGCA	CAGCGGCATC	TACGCTGCG	ATGTCAGCAA	TGAGTTCTCC	TCAAGGGATT	1200
CTCAGGTAC	TGTGGATGTT	CTTGACCCCC	AGGAAGACTC	TGGGAAGCAG	GTGGACCTAG	1260
TGTCAGGCTC	GCTGGTGTG	GTGGGTGTG	TGCGCGCACT	CTTGTGCTGC	CTTCTGGTGG	1320
TGGTGGTGGT	GCTCATGTC	CGATACCATC	GGCGCAAGGC	CCACAGATG	ACCCAGAAAT	1380
ATGAGGAGGA	GCTGACCCCTG	ACCAGGGAGA	ACTCCATCCG	GAGGCTGCAT	TCCCACATACA	1440
CGGACCCAG	GAGCCAGCCG	GAGGAGAGTG	TAGGGCTGAG	AGCCGAGGGC	CACCCCTGATA	1500
GTCTCAAGGA	CAACAGTAGC	TGCTCTGTG	TGAGTGAAGA	GCCCCAGGGC	CGCAGTTACT	1560
CCACGCTGAC	CACCGTGAGG	GAGATAGAAA	CACAGACTGA	ACTGCTGTCT	CCAGGCTCTG	1620
GGCGGGCCGA	GGAGGAGGAA	GATCAGGATG	AAGGCATCAA	ACAGGCCATG	AACCATTG	1680
TTCAGGAGAA	TGGGACCTCA	CGGGCCAAGC	CCACAGGGCAA	TGGCATCTAC	ATCAATGGC	1740
GGGGACACCT	GGTCTGACCC	AGGCTGCGT	CCCTTCCCTA	GGCCTGGCTC	CTTCTGTTGA	1800
CATGGGAGAT	TTTAGCTCAT	CTTGGGGGCC	TCCTTAAACA	CCCCCATTTC	TTGCGGAAGA	1860
TGCTCCCTCAT	CCCACTGACT	GCTTGAACCT	TACCTCCAAAC	CCTTCTGTT	ATCGGGAGGG	1920
CTCCACCAAT	TGAGTCTCTC	CCACCATGCA	TGCAAGTCAC	TGTTGTTG	CATGTGTGCC	1980
TGTTGAGTG	TTGACTGACT	GTGTTGTTG	GGAGGGTGTG	CTGTCCTGG	AGGGGTGACT	2040
GTGTCCTGGG	TGTGTATTAT	GCTGTCATAT	CAGAGTCAG	TGAACGTG	TGTATGTGCC	2100
ACGGGATTTG	AGTGGTTGCG	TGGGCAACAC	TGTCAGGGTT	TGGCGTGTG	GTCATGTGGC	2160
TGTTGTTGAC	CTCTGCTGA	AAAAGCAGGT	ATTTTCTCAG	ACCCCAGAGC	AGTATTAAATG	2220
ATGCCAGAGGT	TGGAGGAGAG	AGGTGGAGAC	TGTGGCTCAG	ACCCAGGTG	GGGGGCATAG	2280
CTGGAGCTGG	AAATCGCTC	CGGTGTGAGG	GAACCTGCTC	CCTTACCACT	CGGAGGCCATG	2340
GGGGCAAGTG	TGAAGCACCC	AGTCCCTGGG	TCAGCCAGAG	GCTTGAAC	TTACAGAAC	2400
CCTCTGCCCT	CTGGTGGCCT	CTGGGCGCTGC	TGCACTGACA	TATTTCTG	AAATATACAT	2460
GCGCCGGAG	CTTCTTGCG	GAATACTGCT	CCGAATCACT	TTTAATTTT	TTCTTTTTTT	2520
TTCTCTGCC	TTTCCATTAG	TTGTATTTT	TATTTATTT	TATTTTTT	TTTTTTT	2580
GATGGAGTCT	CACTATGTT	CTCAGGCTGG	CCTTGAAC	CTGGGCTCAA	GCAATCCTCC	2640
TGCCCTAGCC	TCCCCTAGTAG	CTGGGACTTT	AAAGTGTACAC	CACTGTCCT	GCTTGAATC	2700
CTTTACGAAG	AGAAAAAAA	AATTAAAGAA	AGCCTTTAGA	TTTATCCAAT	GTTTACTACT	2760
GGGATTGCTT	AAAGTGAGGC	CCCTCCAACA	CCAGGGGTT	AATTCTGTG	ATTGTGAAAG	2820
GGGCTACTTC	CAAGGCATCT	TCATGCGAGG	AGCCCCCTGG	GAGGGCACCT	GAGAGCTGGT	2880
AGAGTCTGAA	ATTAGGGATG	TGAGCCTCGT	GGTTACTGAG	TAAGGTA	TTGCATCCAC	2940
CATTGTTG	GATACCTTAG	GGAAATGCTT	GGACCTGTTG	ACAAGGGCTC	CTGTTCAATA	3000
GTGGTGTG	GGAGAGAGAG	AGCAGTGATT	ATAGACCGAG	AGAGTAGGAG	TTGAGGTGAG	3060
GTGAAGGAGG	TGCTGGGGT	GAGAATGTCG	CCTTCTCCCC	TGGGTTTTG	ATCACTAATT	3120
CAAGGCTCTT	CTGGATGTTT	CTCTGGGTG	GGGCTGAGT	TCAATGAGG	TTATTTT	3180
CTGGGCCACCC	CAGATACAT	CAGCCAGAAAT	ACTAGATGTT	AGTACCCAA	CTCTTCTTAG	3240
TCTGAAATCT	GCTGGATTTC	TGGCTTAAGG	GAGAGGCTCC	CATCCTTCG	TCCCCAGCCA	3300
GCCTAGGACT	TCGAATGTGG	AGCCTGAAGA	TCTAAGATCC	TAACATGTAC	ATTTATGTA	3360
AATATGTGCA	TATTGTACA	TAAGATGATA	TTCTGTTTT	AAATAAACAG	ACAAAACTTG	3420
TCAAAAAAA	AAAAAAA	AAAAAAA	AAAAAAA			

BCX5 Protein sequence (SEQ ID NO:6)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Protein Accession #: AF160477; Predicted Signal sequence: 1-26; Predicted TM domains: 355-371; PFAM domains: IgSF_domain: 45-129, 162-225, 263-317; Summary: A type Ia TM protein; is a member of the immunoglobulin superfamily.

MPLSLGAEMW	GPEAWLILL	LLASFTGRCP	AGELETSDVV	TVVLGQDAKL	PCFYRGDSGE	60
QVGQVAWARV	DAGEGAQELA	LLHSKYGLHV	SPAYEGRVEQ	PPPPRNPLDG	SVLLRNAVQA	120
DEGEYECRVS	TFPAGSFQAR	LRLRVMPVPL	PSLNPGPALE	EGQGLTLAAS	CTAEGPSAPS	180
VTWDTEVKGT	TSSRSFKHSR	SAAVTSEFH	VPSRSMNGQP	LTCVVSHPGL	LQDQRITHIL	240
HVSFLAEASV	RGLEDQNWLH	IGREGAMLKC	LSEGQPPPSY	NWTRLDGPPLP	SGVRVDGDTL	300
GFPLLTTEHS	GIYVCHVSN	FSSRDSQVTV	DVLDPQEDSG	KQVDLVSASV	VVVGIAALL	360
FCLLVVVVV	MSRYHRRKAQ	QMTQKYEEEL	TLTRENSIRR	LHSHHTDPRS	QPEESVGLRA	420
EGHPDSLKD	SSCSVMSEEP	EGRSYSTLTT	VREIETQTEL	LSPGSGRAEE	EEDQDEGIKQ	480
AMNHFVQENG	TLRAKPTGNG	IYINGRGLHV				

mouse BCX5 Protein sequence (SEQ ID NO:7)

Gene name: mouse_LNIR; Unigene number: n/a; Probeset Accession #: BF168327; Protein Accession #: n/a; Predicted Signal sequence: 1-27; Predicted TM domains: 346-362; PFAM domains: IgSF_domains:44-126, 166-221, 259-313; Summary: This is the mouse orthologue of human BCX5; it is a type Ia TM protein of unknown function.

MPLSLGAEMW	GPEAWLRLF	LASFTGQYSA	GELETSDVVT	VVLGQDAKLP	CFYRGDPDEQ	60
VGQVAWARVD	PNEXYPGAGL	LHSKYGLHV	PAYEDRVEQX	XHETFRSVL	LRNAVQAEG	120

EYECRVSTFP SGSFQARMRL RVLVPLPSL NPGPPLLEQQ ADVAASCTAE GSPAPSVTWD 180
 TEVKGTQSSR SFTHPRSAAV TSEFHLPVSR SMNGQPLTCV VSHPGLLQDR RITHTLQVAF 240
 LAEASVRGLE DQNLWQVGRE GATLKCLSEG QPPPKNWTR LDGPLPSGVR VKGDTLGFPP 300
 LTTEHSGVYX CHVSNELSSR DSQVTVEVLD PEDPGKQVDL VSASVIIVGV IAALLFCLLV 360
 VVVLMSRYH RRKAQQMTQK YEEELLTRE NSIRRLHSHH SDPRSQPEES VGLRAEGHPD 420
 SLKDNSSCSV MSEEPEGRSY STLTTVREIE QTTELLSPGS GRTEEDDDQD EGIKQAMNHL 480
 CRKMGP

BCZ6 DNA sequence (SEQ ID NO:8)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Nucleic Acid Accession #: NM_002184;
 Coding sequence: 256-3012 (start and stop codons underlined)

GAGCAGCCAA AAGGCCCGCG GAGTCGCCTG GGGCCGCCCG GGCAGCTG AACGGGGGC 60
 CGCGCCTGCC AGGGCGACGG GTCTGCCCA GCCTGGCGCC AAGGGGTTCG TGCGCTGTGG 120
 AGACGCGGAG GGTGAGGGCG GCGGCCCTG AGTGAACCCC AATGGAAAAA GCATGACATT 180
 TAGAAGTAGA AGACTTAGCT TCAAATCCT ACTCCTTCAC TTACTAATTG TGTGATTGG 240
 AAATATCCGC GCAAGATGTT GACGTTGCAG ACTTGGGTAG TGCAAGCCTT GTTTATTTC 300
 CTCACCACTG AATCTACAGG TGAACCTCTA GATCCATGTG GTTATATCAG TCCTGAATCT 360
 CCAGTTGTAC AACATTCATTC TAATTTCACT GCACTTGTG TGCTAAAGGA AAAATGTATG 420
 GATTATTTTC ATGTAATTCG TAATTACATT GTCTGGAAA CAAACACATT TACTATTCC 480
 AAGGAGCAAT ATACTATCAT AAACAGAACCA GCATCCAGTG TCACCTTAC AGATATAAGCT 540
 TCAATTAAATA TTCAGCTCAC TTGCAACATT CTTACATTCTG GACAGCTTGA ACAGAAATGTT 600
 TATGGAATCA CAATAATTTC AGGCTTGCC CTAGAAAAAC CTAAAAATTG GAGTTGCATT 660
 GTGAACGAGG GGAAGAAAAT GAGGTGTGAG TGGGATGGTG GAAGGGAAAC ACACTTGGAG 720
 ACAAACTTC A CTTTAAATC TGAATGGCC ACACACAAGT TTGCTGATTG CAAAGCAAA 780
 CGTGACACCC CCACCTCATC CACTGGTGTAT TATTCTACTG TGTTATTGTG CAAACATTGAA 840
 GTCTGGGTAG AAGCAGAGAA TGCCCCCTGGG AAGGTTACAT CAGATCATAT CAATTTGTAT 900
 CCTGTATATA AAGTGAAGCC CAATCCGCCA CATAATTCTT CAGTGTCA CTCAGAGGAA 960
 CTGTCTAGTA TCTTAAATTG GACATGGACC AACCCAAGTA TTAAAGGTGT TATAATAACTA 1020
 AAATATAACA TTCATATAG GACCAAAAGT GCCTCAACTT GGAGCCAGAT TCCTCCGTAA 1080
 GACACAGCAT CCACCGGATC TTCACTTCACT GTCACAGACC TTAAACCTTT TACAGAATAT 1140
 GTGTTTAGGA TTGCGCTGTG AAGGAAAGAT GGTAGGGAT ACTGGAGTGA CTGGAGTGA 1200
 GAAGCAAGTG GGATCACCTA TGAAGATAGA CCATCTAAAG CACCAAGTTT CTGGTATAAA 1260
 ATAGATCCAT CCCATACTCA AGGCTACAGA ACTGTACAAC TCGTGTGAA GACATTGCCT 1320
 CCTTTTGAAG CCAATGGAAA AATCTTGGAT TATGAAAGTGA CTCTCACAAAG ATGGAATCA 1380
 CATTTACAAA ATTACACAGT TAATGCCCA AACTGACAG TAAATCTCAC AAATGATCGC 1440
 TATCTAGCA CCCTAACAGT AAGAAATCTT GTTGGCAAT CAGATCAGC TGTTTTAACT 1500
 ATCCCTGCCT GTGACTTTCA AGTACTCTAC CCTGTAACTGG ATCTTAAAGC ATTCCCCAAA 1560
 GATAACATGC TTTGGGTGGA ATGGACTACT CCAAGGGAAAT CTGTAAGAA ATATATAACTT 1620
 GAGTGGTGTG TGTTATCAGA TAAAGCACCCT GTGATCACAG ACTGGCAACA AGAAGATGGT 1680
 ACCGTCGATC GCACCTATT AAGAGGAAAC TTAGCAGAGA GCAAAATGCTA TTGATAACA 1740
 GTTACTCCAG TATATGCTGA TGGACCAAGA AGCCCTGAAT CCATAAAAGGC ATACCTTAAA 1800
 CAAGCTCCAC CTTTCAAAG ACCTACTGTG CGGACAAAGA AAGTAGGGAA AAACGAAGCT 1860
 GTCTTAGAGT GGGACCAACT TCCTGGTGTG GTTCAGAATG GATTTATCAG AAATTATAACT 1920
 ATATTTTATA GAACCATCAT TGGAAATGAA ACTGCTGTGA ATGTTGGATT TTCCCACACA 1980
 GAATATACAT TGTCCTCTT GACTGTGAC ACATTTGTACA TGGTACGAAT GGCAGCATAAC 2040
 ACAGATGAAG GTGGGAAGGA TGGTCAGAA TTCACTTTA CTACCCCCAA GTTTGCTCAA 2100
 GGAGAATTG AAGCCATAGT CGTGCCTGTG TGCTTAGCAT TCCTATTGAC AACTCTCTG 2160
 GGAGTGTGTG TCTGCTTAA TAAGCGAGAC CTAATTAAA AACACATCTG GCCTAATGTT 2220
 CCAGATCCTT CAAAGAGTC TATTGCCAG TGTCACACTC ACACCTCTCC AAGGCACAAT 2280
 TTTAATTCAA AAGATCAAAT GTATTCTAGT GGCATTTCCTA CTGATGTAAG TGTTGTGAA 2340
 ATAGAAGCAA ATGACAAAAA GCCTTTCCA GAAGATCTGA AATCATTGGA CCTGTTCAA 2400
 AAGGAAAAAA TTAATCTGA AGGACACAGC AGTGGTATTG GGGGGTCTTC ATGCATGTCA 2460
 TCTCTTAGGC CAAGCATTTC TAGCAGTGTG GAAAATGAAT CTTCACAAA CACTTCGAGC 2520
 ACTGTCCAGT ATTCTACCGT GGTACACAGT GGCTACAGAC ACCAAGTTCC GTCACTCCAA 2580
 GTCTTCTCAA GATCCGAGTC TACCCAGCCC TTGTTAGATT CAGAGGAGCG GCCAGAAGAT 2640
 CTACAAATTAG TAGATCTGT AGATGGCGGT GATGGTATT TGCCCTGGCA ACAGTACTTC 2700
 AAACAGAAACT GCAGTCAGCA TGAATCCAGT CCAGATATTG CACATTGAA AAGGTCAAAG 2760
 CAAGTTTCAT CAGTCATGA CAAAGATTGGT GTTAGACTTA AACAGCAGAT TTCAGATCAT 2820
 ATTCACAAT CCTGTGGATC TGGGCAAATG AAAATGTTTCA AGGAAGTTTC TGCACTGGAT 2880
 GCTTTGGTC CAGGTACTGA GGGACAAGTA GAAAGATTG AAACAGTTG CATGGAGGCT 2940
 GCGACTGATG AAGGCATGCC TAAAAGTTAC TTACCAACAGA CTGTACGGCA AGGCAGCTAC 3000
 ATGCCTCAGT GAAGGACTAG TAGTTCTGC TACAACCTCA GCAGTACCTA TAAAGTAAAG 3060
 CTAAAATGAT TTTATCTGTG AATTC

BCZ6 Protein sequence (SEQ ID NO:9)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Protein Accession #: NP_002175; Predicted
 Signal sequence: 1-22; Predicted TM domains: 625-641; PFAM domains:
 fibronectin_type_III_domains: 222-311, 424-509, 519-606; Summary: A type I TM protein; it
 homodimerizes or heterodimerized to make a functional receptor for IL-6, oncostatin-M, IL-11,
 LIF, and CNTF.

MLTLQTWWVQ ALFIFLTTE S TGELLDPCGY ISPESPVQL HSNFTAVCVL KEKCMDYFHV 60
 NANYIVWKTN HFTIPKEQYT IINRTASSVT FTDIASLNIQ LTCNILTGFQ LEQNIVYGITI 120

I	SGLPPEKPK	NLSCIVNEGK	KMRCEWDGGR	E	THILETNFTL	KSEWATHKFA	DCKAKRDTPT	180
S	C	TVDYSTVY	FVNIEVWVEA	E	NALGKVTS	HINFDPVYKV	KPNPPHNL	240
L	T	WTNPSIK	SVIILKYN	I	YRTKDA	STWS	QIPPE	300
K	M	EDGKGW	SDWSEEAS	G	TYEDRPSK	AP	SSFTVQDLK	360
G	A	ILDYEVTL	TRWKSHLQ	N	TVNATKLTVN	L	TQGYRTVQLV	420
F	O	ATHPVMDL	KAFPKDNM	LW	LTNDRYLATL	T	WKTLPF	480
Y	L	RGNLAESK	VEWTPRESV	KKYI	LEWCVL	S	AAVL	540
Q	P	LPDVQNGF	IRNYTIFYR	T	ADPGSPESI	K	KGPTV	600
K	D	GPEFTFT	PKFAQGEIEA	I	NETAVNV	S	RTKKV	660
S	H	IAQWSPHT	PPRHNFNSKD	Q	DSDGNFTD	V	GKNEAV	720
T	E	GHSSGGIGG	SSCMSSSRP	I	SSDEN	Q	LEWD	780
E	S	TOPPLLDSE	ERPEDLQLVD	Y	WTTPRESV	T	QLEDGT	840
N	E	EEFDVRLQ	QISDHISQSC	C	ADPGSPESI	K	TVHSGYRHQ	900
E	D	MPKSYLPQ	GSGQM	P	KAYLKQAPP	S	VPSVQVF	
E	Y	TVQYMPQ	MFQE	V	QHESSPDISH	F	SRS	
Q	Y	NEEDFVRLQ	VSAADAFGPG	T	TEGQVERFET	V	FERSKQVSS	
R	Y	QISDHISQSC	TEGQVERFET	V	VGMEAATDEG			
Q	Y	MPKSYLPQ	QGQYMPQ					

BF4 DNA sequence (SEQ ID NO:10)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
Nucleic Acid Accession #: AB020689; Coding sequence: 108-2777 (start and stop codons underlined)

GAAC	TTATGT	AGCCTCATTA	TCCC	GCTCCG	TGAGGTGACA	ATTGTG	GGAAA	AGGCAGACAG	60		
CTCC	AGTG	TG	CTCCCAGTC	CC	TATCAC	TCAGC	ACCCG	AAACAGGATG	120		
TTGC	CAACTT	GAAAGATAGA	GACT	TTCTAG	TGCA	GAGGAT	CTCAG	ATTT	CTGCAACAGA	180	
CTACTT	CCAA	AATATATTCT	GACA	AGGGAGT	TTG	CAGGAAG	TT	ACAACAGT	TCAGATGATG	240	
AGGT	GACT	TCGACCCAGC	AGC	CTCGTCT	CCT	CCGCCCC	CC	AGAGAACG	ACGAGCTCTG	300	
ATGCTG	ATGG	AGAGGCCAG	TTT	AACTTA	ATG	CCAACAC	CGT	CCCACACACA	GCCACACAGA	360	
CCTG	TATG	GAC	CATG	TATCGG	CGG	GCTCTC	CCG	GAGGAGTT	CAACCCGAAA	420	
AGTTT	CTGAA	AGAGCAAGCC	TGG	AAAGATT	CT	TTGCTG	TA	TGCA	GGGATCTGCA	480	
TGTACCG	CAC	AGAGAAAACG	CGG	GACTG	TTG	TGAAGGG	CAT	CCC	GGAG	540	
GGGAGC	CTCTG	GCTG	TCAGG	TGCCA	TCA	ATGAGAA	GG	CCACACAT	CCTGGTACT	600	
ATGAAGAC	CT	AGTGGAGAAG	TCC	ATGGGG	AGT	TATAATCT	CG	CCACGGAG	GAGATTGAGA	660	
GGGATT	TACA	CCG	CTTCT	CC	CAAGAACACC	CAG	CTTT	TC	GAATGAAATG	GGCATTGCTG	720
CACTA	AGGG	AGT	CTTAA	ACA	GCTT	ATG	AA	TC	TATTGCCAGG	780	
CCATG	AATAT	TGT	CACTTCA	GTG	CTGCTG	CTT	ATG	CCAA	GAGGAGGAA	GCTTCTGGC	840
TGCTT	GTCG	TTT	GTG	TGAG	CG	GATTACTA	CA	ACACCAGA	GTTG	TGGGTG	900
CACTGG	TGGA	CCA	AGGTGTC	TTT	GAGGAGG	TAG	CAC	GAGA	CTACGTCCC	CAGCTGACG	960
ACTG	CATG	CA	AGAC	TGGG	GTG	ATTCTCA	CC	ATCT	CTTC	CTCACACTAT	1020
TTC	CAGT	G	TG	ATG	CTT	CC	TC	TC	TC	TGACTGTTTC	1080
GAATT	AAAGT	GAT	ATT	CCAG	TTG	CC	CTG	AA	TG	TTCATGAAAG	1140
TGA	ACTG	CAA	GGAT	GATGGG	GAG	GGG	CTG	AA	GT	GACAGTGTGA	1200
CCAA	TAAAGA	AG	CAC	ACATG	CCT	CC	TC	AC	CT	AGCGATGATG	1260
TGGA	AACTT	CC	CTG	GAGG	TA	CA	CTC	AC	CT	GAATGAAATG	1320
GAAC	ATATCCG	GGC	CAGATT	TTG	GAACAGA	TG	AGATT	CA	AC	AGAGACTG	1380
AGAC	GCTG	GA	GA	TACT	AA	ACG	CAAC	TG	AG	AAAGTGTAC	1440
TTAC	CATTG	TG	AG	CTG	TTATG	CT	CTT	TT	CA	CTCACCACT	1500
GCT	ACTGGG	GG	GG	GAGC	AA	CGC	GCTG	CC	CC	CCCCGCTG	1560
AA	ACAGT	ATG	CAT	G	AG	CA	CTG	AC	CT	TGCTCTCTC	1620
CATG	GGG	AC	TG	CTG	GG	CT	CTG	TT	TT	TTTCTTGGG	1680
GAGA	ACT	CTT	GAT	TTG	GG	CTT	GGG	TT	AT	CCATGGGG	1740
AC	CTC	AC	AGA	AA	CTC	AC	CTC	AC	CT	CCATCTCTG	1800
AT	CAAG	ATG	TA	CTC	TTG	TA	CTC	TT	TT	GATATTACCC	1860
CA	AA	ATG	TG	TTG	GG	TTG	GG	TT	TT	GAACAGTAC	1920
TTG	TACGG	GG	GG	ATTG	GG	GG	GG	GG	GG	GGGATGGCT	1980
ATT	TTT	GAG	AG	CTG	GACT	CC	GG	GG	GG	GGGAGGGG	2040
CCA	AA	TTAA	TC	AGGG	GG	GG	GG	GG	GG	GGGAGGGG	2100
AAG	ACCC	AA	TG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2160
AGA	TGGG	GG	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2220
GCAG	TGGG	GG	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2280
GCC	AGC	CTT	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2340
AGG	AA	ACT	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2400
ACA	AC	GGG	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2460
TGT	CCT	CA	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTGCTG	2520
AA	AT	CGG	GAG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2580
GC	AC	AGG	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2640
CTG	AGC	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	2700
GTG	AAAAAA	CAT	CCGG	AT	GGG	CAAG	GG	GG	GG	GGGAGGGG	2760
CGG	CAT	G	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2820
GAT	TTT	TAT	GT	TCT	CTG	TG	GG	TT	TT	GGGAGGGG	2880
TA	GT	AC	TTG	GG	GG	GG	GG	GG	GG	GGGAGGGG	2940
AA	AT	TTT	AA	AC	CTG	GG	GG	GG	GG	GGGAGGGG	3000
GA	AT	AC	GG	GG	GG	GG	GG	GG	GG	GGGAGGGG	3060
TGT	ATT	AA	CTT	AA	TC	AA	TC	AA	TC	GGGAGGGG	3120
TTA	ATG	CTA	TTG	CTG	TG	GCT	GT	GG	GT	GGGAGGGG	3180
CTG	TA	AG	TA	GG	GG	GG	GG	GG	GG	GGGAGGGG	3240
CG	AGG	GG	AA	ATG	AA	AC	GG	GG	GG	GGGAGGGG	3300
AT	TT	CAT	AG	TG	CT	AC	CT	AA	TC	GGGAGGGG	3360

ACATTTGGG	TTCTAAACAG	TTATGAAAGT	AAGTGATT	TACATGATT	CAGAATAACA	3420					
CTTGATTA	CCAATTTAGA	CAGATACCAG	ACCAATT	CATTTAAGAA	ATTGTTCTGA	3480					
TTATTTACGT	CAACTCATTA	GAATTCACTG	AAAAGTAACA	GTC	CTTGC	3540					
TGAAAGTAGC	AGCAAAGACA	GAGGGCTCAT	GACAGGTTT	TGCTTTGCT	TTGCTTTGCT	3600					
TTTGAAAGA	GTAAAAGTAA	TGATGCTTCT	GATACTGGAT	GTTAGCTTC	TTACTGCAA	3660					
AACATAAGT	AAACAGTC	CTTACCAT	CCCGTATTCT	CCATAGATTG	AAGAAATT	3720					
TACCA	CATGAC	CACATTTCCA	TCAAATCAAT	GTAGAGATAA	TGTAAACTGA	3780					
AAAAAAATCT	GCAAGATAAT	GTA	ACTGAAT	GT	TTTAAAAA	3840					
AAAAGAATAG	TATGCTCTAT	TTCC	TGAATG	GATG	TGAAAAGCTAG	3900					
CTTGAAATT	TTGCTTCTT	TTTATTACTG	TTATGATT	GCT	TTTAC	3960					
GATTTTCT	TCTGATTG	GAATT	CATA	TCATGGT	CTTCTTGG	4020					
ATATTTC	CAACACATTC	CTT	TTTTA	TTATACATTG	TGTC	CTT	TTG	4080			
CTGCTGTTG	TTTTTATTCT	ATT	TACAGGA	TGAT	TTTAA	ACTG	CAAAT	GAAG	TAGTGT	4140	
TAACCTAAA	TAGGCTAAAT	GTG	AACAAAT	AAA	ATACAGC	AA	AA	ACTCAG	AAAA	AAAA	4200
AAAAAA	AAAA										

BFG4 Protein sequence (SEQ ID NO:11)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
 Protein Accession #: BAA74905; Signal sequence: none; Predicted TM domains: 302-318; PFAM
 domains: TBC_domain: 135-347; Summary: a Type II membrane protein, likely localized to the
 peroxisome.

MTFLFANLKD	RDFLVQRISD	FLQQTTSK	IY	SDKEFAGSY	N	SSDDEVYSRP	SSLVSSSPQR	60
STSSDADGER	QFNLNNGNSVP	TATQTLMTMY	RRRSPEEFNP	KLAKEFLKEQ	AWKIHFAEYG	120		
QGICMYRTEK	TRELVLKGIP	ESMRGELWLL	LSGAINEKAT	HPGYYEDLVE	KSMGKYNLAT	180		
EEIERDLHRS	LPEHMPAFQNE	MGIAALRRLV	TAYAFRNPNI	GYCQAMNIVT	SVLLLYAKEE	240		
EAFWLLVALC	ERMLPDPYNT	RUVGALVDQG	VFEELARDYV	PQLYDCMQLD	GVISTISLSW	300		
FLTFLLSVM	FESAVVVVDC	FFYEGIKVIF	QLALAVLDAN	VDKLLNCKDD	GEAMTVILGRY	360		
LDSVTNKDST	LPPIPHHLHSL	LSDDVEPYPE	VDI	FRLIRTS	YEKFGTIRAD	420		
LKVIQTLED	TKRNVVRTIV	TETSFTIDEL	EELYALFKAE	HLTSCYWGG	SNALDRHDP	480		
LPYLEQYRID	FEQFKGMFAL	LFPWACGTHS	DVLASRFLQ	LDENGDSLIN	FREFVSGLSA	540		
ACHGDLTEKL	KLLYKMHVLP	EPSSDQDEPD	SAFEATQYFF	EDITPECTHV	VGLDSRSKQG	600		
ADDGFVTVSL	KPDGKCRANS	QENRNLYRLW	TPENPKSKSN	AKDLPKLNQG	QFIELCKTMY	660		
NMFSEDPNEQ	ELYHAAAVT	SLLLEIGEVG	KLFVAQPAKE	GGSGGSGPSC	HQGIPGVLF	720		
KKPGQPYVV	ESVEPLPASL	APDSEEHSLG	GQMEDIKLED	SSPRDNGACS	SMLISDDDTK	780		
DDSSMSSY	LSAGSHEDK	LHCEEIGEDT	VLVRSGQGTA	ALPRSTSLDR	DWAITFEQFL	840		
ASLLTEPALV	KYFDKPVCM	MMGKPLTSAS	DYEISAMSG					

BCU7 DNA sequence (SEQ ID NO:12)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Nucleic Acid
 Accession #: n/a; Coding sequence: 1-573 (stop codon underlined)

TATTTTATTT	TCCAGGCTAA	AGCAAATGAA	AGTTGCTGG	TATCAACACA	GCCTGCCATA	60		
TTTTTCACAG	CATGCAACAA	TGGTGTCTAGG	ATAGCTATT	CTTACTGTAA	TTGCCAGAGG	120		
CAGAAATGGT	CTGGGTATAA	GCTATT	CAT	AAAAGCAGCT	TTAAATTGTC	180		
TTTCATGTG	GAAAGGTGTC	ATT	CAAAA	AAAGTAATTG	GCATACATAT	240		
CGATCCCTC	TGTGGTGT	TTT	TTTTT	ATGACCA	GAAAAATT	300		
CAATATAGGT	TTGGGGCTT	CCATATC	ATC	AAAAGACTG	AA	ATT	TTCTCA	360
AAACTGATGG	ATTCATTT	AGAAT	ATC	TGAGTGTG	TAGACACAGT	420		
CTGGTTATGA	CAGATAAGT	TGCT	CAAAA	ATGTGGATG	AGCCATT	480		
TTATTGCTTC	TGTTCA	GTG	TAT	TATC	TGGCCC	540		
TTGCCCTACA	AATT	TCAT	TTCT	CTG	ATCA	600		
TTCATTTGAA	AGT	TCAT	GC	AG	AACT	660		
ATT	TTACAG	AAAT	TTGAA	TTGAA	CTC	720		
GAAACCAACT	TATCTGCATA	AT	AAAT	CTA	AT	780		
TGTTGATGCC	ATGCTTATCA	AATACATGCA	CAAGCTAAC	ATAATTGAA	TGGCTCTATG	840		
AAGGAAAAT	AATGCTT	AGA	CTTGGTGT	GGT	CTTCC	900		
TCTGCA	CGAAGGATG	AAAT	GGT	GAC	GTC	960		
TTAGGAAGGA	TAAAGT	TACA	TTGCTTAA	TTCTA	AC	1020		
ATAAAAAATT	GTCATT	ATT	GTG	CTAA	AC	1080		
CATGCAAAT	AATT	TTAA	ATT	GTG	AT	1140		
AAAAACTTAG	AAGGTTAAC	ACT	CACTG	TA	AC	1200		
CACATATCCT	TGCTATA	AGT	GGT	GG	CC	1260		
GAAC	GTG	TAG	TAT	CTG	TT	1320		
TGGCTT	ATGCTC	TT	ATC	GGT	AT	1380		
ATT	CA	AA	TTG	TT	CTG	1440		
ACTCTGTTGT	TTAATT	CC	ATG	TT	AC	1500		
TTC	TTG	AA	AGT	TC	AG	1560		
ATATCAC	AACTGG	TTT	TTA	TC	GT	1620		
TCTGGT	ATGAAG	TTT	TTA	CT	GG	1680		
GTAC	CA	TT	CTG	AC	GG	1740		
ACCTAAAATG	TTTGATA	TTA	ACATG	TCA	AA	1800		
GGCACCCACT	TTCTCTAG	TTG	TTCC	TTG	GG	1860		
TGCTT	TTCTG	TTA	TTG	TTG	GG	1920		
TTT	TTCTG	TTA	TTG	TTG	GG	1980		
ATGTGATCATT	TATT	TTA	TTG	TTG	GG	2040		

TTTCATATAT	AAAAGGATTA	TTTCTCCACC	TTTAATTATT	GGCCTATCAT	TTGTTAGTGT	2100
TATTTGGTCA	TATTATTGAA	CTAATGTATT	ATTCATTCA	AAGTCTTCT	AGATTTAAA	2160
ATGTATGCAA	AAGCTTAGGA	TTATATCATG	TGTAACTATT	ATAGATAACA	TCCTAACCT	2220
TCAGTTAGA	TATATAATTG	ACTGGGTGTA	ATCTCTTTG	TAATCTGTTT	TGACAGATTT	2280
CTTAAATTAT	GTTAGCATAA	TCAAGGAGA	TTTACCTTGAA	AGCACTTTC	AAATTGATAC	2340
TTTCAAACCTT	ATTTAAAGC	AGTAGAACCT	TTTCTATGAA	CTAAATCACA	TGCAAACACT	2400
CAACCTGTAG	TATACATAAA	ATGAACTAC	TTATTCTCT	CACCTCTCC	AGTGCCTAGG	2460
AATATTCTTC	TCTGAGCCT	AGGATTGATT	CTATCACACA	GAGCAACATT	AATCTAAATG	2520
GTTTAGCTCC	CTCTTTTTC	TCTAAAAACA	ATCAGCTAAT	AAAAAAAAAA	TTTGAGGGCC	2580
TAATTATTAT	CAATGGTTG	TTGAATATT	CAGTCAGTT	TGTACCTGTT	AGCAGTCTTT	2640
CAGTTGGGG	GAGAATTAAA	TACTGTGTA	AGCTGGTGCT	TGGATACATA	TTACAGCATC	2700
TTGTGTTTA	TTTGACAAAC	AGAATTGGG	TGCGATAATA	TTTTGAGAAT	TAGAGAAGAT	2760
TGTGATGCAT	ATATATAAAC	ACTATTTTA	AAAAATATCT	AAATATGTCT	CACATATTAA	2820
TATAATCCTC	AAATATACTG	TACCATTTA	GATATTTTT	AAACAGATTA	ATTTGGAGAA	2880
GTTTATTCA	TTACCTAATT	CTGTGCCAA	AATGGTGCT	CTGATGTTG	GATATAGTAT	2940
TGTCAGTGTG	TACATATATA	AAACCTGTGT	AAACCTCTGT	CCTTATGAAC	CATAACAAAT	3000
GTAGTTTTT	AAAGTCATT	GTATTGTTT	TTCTTCAAT	AAAAGAGTAT	AATTAATTGG	3060
TTGTTTTGAA						

BCU7 Protein sequence (SEQ ID NO:13)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 125-141, 154-170; PFAM domains: none; Summary: A type III membrane protein, highly overexpressed in breast cancer and prostate cancer; unknown function.

YFIFQAKANE	SLLVSTQPAI	FFTACNNGAR	IAISYCNQR	QKWSGYKLFH	KSSFKLSVLR	60
FSCGKVSFKK	KVIGIHIPHH	RSSLWCXFFY	MTSRKILIFS	QYRFWFHFII	KRLKNYNFRI	120
KLMDFIIELS	VSCVDTQLMF	LVMTDKFAQK	MWMKPLLLLL	LLLLFSCLSI	IPSVAHAAE	180
LPYKFHLAAP						

BFA1 DNA sequence (SEQ ID NO:14)

Gene name: calsyntenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Nucleic Acid Accession #: NM_022131; Coding sequence: 11-2878 (start and stop codons underlined)

TGCTGCGAGG	ATGCTGCCTG	GGCGGCTGTG	CTGGGTGCCG	CTCCTGCTGG	CGCTGGCGT	60
GGGGAGCGGC	AGCGGCGGTG	CGCGGACAG	CCGGCAGCGC	CGCCTCCTCG	CGGCTAAAGT	120
CAATAAGCAC	AAGCCATGGA	TCGAGACTTC	ATATCATGGA	GTCATAACTG	AGAACATGA	180
CACAGTATT	TTGGACCCAC	CACTGGTAGC	CCTGGATAAA	GATGCACCGG	TTCTTTGTC	240
AGGGGAAATC	TGTGCGTTC	AGATCCATGG	CCAGGAGCTG	CCCTTTGAGG	CTGTGGTGCT	300
CAACAAGACA	TCAGGAGAGG	GGCGCTCCG	TGCCAAGAGC	CCCATTGACT	GTGAGTTGCA	360
GAAGGAGTAC	ACATTATCA	TCCAGGCC	TGACTGTGTT	GCTGGGCC	ACGAGACAGC	420
CTGGAAAAAG	TCACACAAAG	CCGTGGTCTA	TATACAGGTG	AAGGATGTC	ACGACTTTGC	480
TCCCACCTTC	AAAGAGCCAG	CCTACAAAGG	TGTTGTCAGC	GAGGGCAAGA	TCTATGACAG	540
CATTCTGCAG	GTGGAGGCC	TGACCGAGGA	CTGCTCCCCA	CAGTACAGCC	AGATCTGCAA	600
CTATGAAATC	GTCACACAG	ATGTGCTTT	TGCCATCGAC	AGAAATGGCA	ACATCAGGA	660
CACTGAGAAG	CTGAGCTATG	ACAACAAACA	CCAGTATGAG	ATCCTGGTGA	CCGCCTACGA	720
CTGTGGACAG	AAGCCCGCTG	CTCACGACAC	CTCTGGTCAG	GTGGATGTC	AGCCAGTTG	780
CAAGCCTGGC	TGGCAAGACT	GGACCAAGAG	GATTGAGTAC	CAGCTGGCT	CCGGGAGCAT	840
GCCCCCTGTC	CCCAGCATCC	ACCTGGAGAC	GTGCGATGGA	GCCCTGTCTT	CCCTCCAGAT	900
CGTCACAGAG	CTGCAAGACT	ATTACATTGG	GAAGGGTTGT	GACCGGGAGA	CCTACTCTGA	960
GAAATCCCTT	CAGAAATTAT	GTGGAGCCTC	CTCTGGCATC	ATTGACCTCT	TGCCATCCCC	1020
TAGCGCTGCC	ACCAAATCGGA	CTGCAAGACT	GCTGGTGGAC	AGCAGTGGAGA	TGATCTTCAA	1080
GTTTGACGGC	AGGCAGGGTG	CCAAAATCCC	CGATGGGATT	GTGCCCAAGA	ACCTGACCGA	1140
TCAGTTCACC	ATCACCATTG	GGATGAAACA	CGGGCTTCCAGC	CCTGGTGTGA	GAGCCGAGAA	1200
GGAAACCATC	CTCTGCACACT	CAGACAAAC	CGAAATGAAC	CGGCATCACT	ATGCCCTGTA	1260
TGTGCACAAAC	TGCCGCTCTG	TCTTCTCTT	GGCGAAGGAC	TTCGACCAAG	CTGACACCTT	1320
TCGCCCGCG	GAGTTCCACT	GGAAAGCTGGA	TCAGATTGTT	GACAAAGAGT	GGCACTACTA	1380
TGTCATCAAT	GTGGAGTTTC	CTGTTGTAAC	CTTATACATG	GATGGAGCAA	CATATGAACC	1440
ATACCTGGTG	ACCAACAGCT	GGCCCATTC	TCCATCTC	ATAGCCATGC	AACTCACAGT	1500
CGGCGCTTGT	TGGCAAGGAG	GAGAAGTCAC	CAAACACAG	TTTGCTCAGT	TCTTCATGG	1560
AAGCTGGCC	AGTCTCACCA	TCCGCCCTGG	CAAATGGAA	AGCCAGAAG	TGATCTCCTG	1620
CCTGCAGGCC	TGCAAGGAAG	GGCTGGACAT	TAATTCCTTG	GAAAGCTTC	GCCAAGGAAT	1680
AAAGTATCAC	TTCAACCCCT	CGCATGCTCAT	CTCTGGTATG	GAAGGGTGAC	ACATTGGAA	1740
CATTAACCGT	GCTCTCCAGA	AAGTCCTCA	CATCACTCC	AGGCAGTTCC	CAACGGCGGG	1800
TGTGCGGGCG	CTCAAAAGT	CCTCCAAAGT	CCAGTGTCTT	GGGGAAAGAC	TATGCATCAG	1860
TATCCCTGAG	GTAGATGCCT	ATGTGATGGT	CCTCCAGGCC	ATCGAGCCCC	GGATCACCC	1920
CGGGGGCACA	GACCACCTCT	GGAGACCTGC	TGCCCAGTT	GAAAGTGCCA	GGGGAGTGAC	1980
CCTCTTCCCT	GATATCAAGA	TTGTGAGCAC	CTTCGCCCCA	ACCGAAGCCC	CGGGGGACGT	2040
GAAAACCAACA	GACCCCAAAAT	CAGAAGTCTT	AGAGGAAATG	CTTCATACT	TAGATTCTG	2100
TGACATTGTTG	GTGATCGGAG	GGGACTTGGGA	CCCAAGGAG	GAGTGCTTG	AGCTCAACCA	2160
CAGTGAAGCTC	CACCAACGAC	ACCTGGATGC	CACTAATTCT	ACTGCAGGCT	ACTCCATCTA	2220
CGGTGTGGC	TCCATGAGCC	GCTATGAGCA	GGTGCTACAT	CACATCCGCT	ACCGCAACTG	2280
GCGTCCGGCT	TCCCTTGAGG	CCCCCGCTTT	CCGGATTAAG	TGCTCAGAAC	TCAATGGCG	2340
CTACACTAGC	AATGAGTTCA	ACTTGGAGGT	CAGCATCCTT	CATGAAGAAC	AACTCTCAGA	2400
TAAGGAGCAT	GTCAATCATC	TGATGTTGCA	GGCTCCCTTC	CTCCAGTCTG	TCCATCATCC	2460
TGAGTCCCCG	ACTAGCATCC	AGCACAGTTC	AGTGGTCCA	AGCATTGCCA	CAGTGGTCAT	2520

CATCATCTCC	GTGTGCATGC	TTGTGTTTGT	CGTGGCCATG	GGTGTGTACC	GGGTCCGGAT	2580	
CGCCCACCA	GA	CTTCATCC	AGGAGACTGA	GGCTGCCAAG	GAATCTGAGA	TGGACTGGGA	2640
CGATTCTGCG	CTGACTATCA	CAGTCACCCC	CATGGAGAAA	CATGAAGGAC	CAGGGCATGG	2700	
GGAAGATGAG	ACTGAGGGAG	AAGAGGAGGA	AGAACGCCAG	GAAGAAATGA	GCTCCAGCAG	2760	
TGCTCTGAC	GACAGCGAAG	AGCAGGAGGA	GGAGGAAGGG	ATGGGCAGAG	GCAGACATGG	2820	
GCAGAATGGA	GCCAGGCAAG	CCCAGCTGGA	GTGGGATGAC	TCCACCCCTCC	CCTACT <u>AGTG</u>	2880	
CCCAGGGTC	TGCTGCCTGG	CCCCATGTC	CCTTTGTAA	ACCTGTACCC	AGTGTATGCC	2940	
CATGTCTATC	ATACCTCAC	TCTGATGTCT	GTGACATGTC	TGGGAAGGCC	TTCTCCAGCT	3000	
TCCGGAGCC	CACCCTTAA	GCCTGGGCA	CTCCCTGTGT	TTCATCCATG	GGGAAGTTC	3060	
AAGAAGGCCA	GCATGGCCAT	CAGTGAGGAC	TTCAGGGTAG	ACTTTGTCCT	GTAGCCTCCA	3120	
CTCTGCTCT	AAGTCCCCA	GCATCTGAC	TACCTGTCTG	CAGAGTTTC	CTTGTGTTT	3180	
TCTGCAAGG	AAGAAGGCC	ACCTTTGTGT	CACTCACCTC	CCCAGGCTCA	GAGTCCCAA	3240	
GGCCCTGGGG	TTCCAACCTA	CTGTGCGTCT	CCTCCACACA	GACCACTAGG	TTCTCCATG	3300	
CTGACTCCAG	GTTGCTCAT	ACAAGGAGGG	TGGTTGAAC	TCACACACGT	AAGGTCTTAG	3360	
TGCTTAACAG	TTTAAAGGAA	AGTCTCTGTG	GAGCAGAAC	TAAGTTTACA	GGGAAAGGTA	3420	
CACACATTCT	CTCTCTCT	CTCTCTGTG	CTATCTAGTT	CCCCAGCTTG	GAGAGCCTTT	3480	
CCCCTTGCTT	CTTCTGAGG	CCATATAAGC	TTATAAGAAA	AGTCCCAAAC	CAAGAATAGG	3540	
TCTCTGGCCA	CAAGCAGGGT	CTGATCCCC	ATCAGAGCTA	TCTGAGCCTG	CCTGTCTGGG	3600	
CACCTGCTGC	AACCATGCAG	CTACCCGCC	AGGGCACTC	AGCAAAACAGA	ACCACAGGGC	3660	
CCAGGAGGCA	TTCCACACAG	GCACTGCC	AGGACAAACAC	AACAAGGACA	GTCAACACAA	3720	
GGACAACAAG	GACACAACAC	AACACACAA	AAGGACAGTC	AACAAACAGC	TAGAGCCAGA	3780	
AAGCAGATGG	AAATGCTAAT	GAGGTCACAC	GTAGGCTCTA	TGCTGGGTTG	AGTGGGGGTG	3840	
GCTGGGCTCC	CCACCGACAG	AGGGGACCT	GAGGTTGGCA	AGGCTCTCAC	CACTCAGCCT	3900	
TATGGTCCCT	TATCTCCTAT	CTTCCCTCTT	GAGAAAATAC	ACGCTTCTG	CATGTATTAG	3960	
AAACGCACGA	GCTCCACCAA	GTCTACAATG	AAAGTTTGAA	ATTTAACTGC	AAGGAATTAG	4020	
AAGCATAATT	GAATCATGT	CAGCTCTTC	TTTCTCTGTC	TCATAAAAAGG	AGGAACACTT	4080	
TAGATAGAGG	GCAAAATATAT	CTGAAAACCT	AATTCTTTTC	TTTTTTGAT	AAGGAAATCT	4140	
TTTCATCTC	CATCTAACAA	TGCAACACCT	GTGAGAGGAA	TTCTTCTAT	AGTAACCTG	4200	
CTGTGATCTT	TTGTGGCCA	GAGAATAGCA	GGCAAGAATT	AGGGCCTG	CAGAATTCC	4260	
ACGAAGCTCT	GAGAACATGT	TTGTTTCGAA	TGCTGTGATTC	CTCTTGTCA	TCAATGTGTA	4320	
TGCTCTGTCC	CCATCTCTCA	CTCCTCTCA	AGTCACACC	AAATTGTTTG	GCACAGGCAC	4380	
AGAGCTGGTC	CCTAGTTAAG	TGGCATTAT	GTTAAAAAAA	A			

BFA1 Protein sequence (SEQ ID NO:15)

Gene name: calsyntenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Protein Accession #: NP_071414; Predicted Signal sequence: 1-20; Predicted TM domains: 832-848; PFAM domains: cadherin_domains: 48-151, 165-254; Summary: A type I membrane protein; a member of the calsyntenin family; is related to the FAT tumor suppressor; is likely an adhesion molecule important in mammalian developmental processes and cell communication.

MLPGRLCWVP	LLLALGVGSG	SGGGGDSRQR	RLLIAAKVNKH	KPWIETSYHG	VITENNDTVI	60
LDPPFLVALDK	DAPVPFAGEI	CAFHKHQEL	PFEAVVNLKT	SGEGRRLAKS	PIDCELQKEY	120
TFIIQAYDCG	AGPHETAWKK	SHKAVVHIVQ	KDVNEFAPTF	KEPAYKAVVT	EGKIYDSILO	180
VEAIDEDSCP	QYSQICNYEI	VTTDVPFAID	RNGNIRNTKE	LSYDKQHQYE	ILVTAYDCGQ	240
KPAQQDTLVQ	VDVKPVCKPG	WQDWTKRIEY	QPGSGSMPLF	PSIHLETCDG	AVSSLQIVTE	300
LQTNYIGKGC	DRETYSEKSL	QKLCGASSGI	IDLLPSPSAA	TNWTAGLLVD	SSEMIFKFDG	360
RQGAKIPDGI	VPKNLTDQFT	ITMWMKHGPS	PVGRAEKETI	LCNSDKTEMN	RHHYALYVHN	420
CRLVFLLRKD	FDQADTFRPA	EHWKLHDQIC	DKEWHYVYVIN	VEFPVVTLYM	DGATYEPYLV	480
TNDWPVHPHS	IAMQLTVGAC	WQGGEVTKPQ	FAQFFHGSIA	SLTIRPGKME	SQKVISCLQA	540
CKEGLDINSL	ESLGQQGIKYH	FNPSQSILVM	EGDDIGNINR	ALQKVSYINS	RQFPTAGVRR	600
LKVSSKVQCF	GEDVCISIPE	VDAYVMVLQA	IEPRITLRTG	DHFWRPAAQF	ESARGVTLFP	660
DIKIVSTFAK	TEAPGDVKTT	DPKSEVLEEM	LHNLDLCDIL	VIIGGDLDPHQ	ECLELNHSEL	720
HQRHLDATNS	TAGYSIYGVG	SMSRYEQVLH	HIRYRNWRPA	SLEARFRFIK	CSELNGRYTS	780
NEFNLEVSIL	HEDQVSDEKHE	VNHLLVQPPF	LQSVHHPESR	SSIQHSSVVP	SIATVIIIS	840
VCMLVFVVA	GYVYRVRIAHQ	HFIQETEAAK	ESEMDWDDSA	LTITVNPMEK	HEPGHGEDE	900
TEGEEEEEAE	EEMSSSSGSD	DSEEEEEE	MGRGRHGQNG	ARQAQLEWDD	STLPY	

BFG7 DNA sequence (SEQ ID NO:16)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Nucleic Acid Accession #: n/a; Coding sequence: <1-906 (stop codon underlined)

CGGGTCGACC	CACCGCGTCCG	GGGAGAAAGG	ATGGCCGGCC	TGGCGGCGCG	GTTGGCTCTG	60
CTAGCTGGGG	CAGCGGGCCT	GGCGAGCGGC	TCCCAGGGCG	ACCGTGAGCC	GGTGTACCGC	120
GACTGCGTAC	TGCACTGCGA	AGAGCAGAAC	TGCTCTGGGG	GCGCTCTGAA	TCACTTCCGC	180
TCCCGCCAGC	CAATCTACAT	GAGCTCTAGCA	GGCTGGACCT	GTCGGGACGA	CTGTAAGTAT	240
GAGTGTATGT	GGGTACCGT	TGGGCTCTAC	CTCCAGGAAG	GTCACAAAGT	GCCTCAGTTC	300
CATGGCAAGT	GGCCCTTCTC	CCGGTTCTG	TTCTTTCAAG	AGCCGGCATC	GGCCGTGGCC	360
TCGTTTCTCA	ATGGCCTGGC	CAGCCTGGT	ATGCTCTGCC	GCTACCGCAC	CTTCGTGCCA	420
GCCTCCTCCC	CCATGTACCA	CACCTGTGTG	GCCTCGGCC	GGGTGTCCCT	CAATGCATGG	480
TTCTGGTCCA	CAGTYTTCCA	CACCAAGGGAC	ACTGACCTCA	CAGAGAAAAT	GGACTACTTC	540
TGTGCCTCCA	CTGTCATCT	ACACTCAATC	TACCTGTGCT	GCGTCAGCCT	CATCCGCTTC	600
GACTATGGCT	ACAACCTGGT	GGCCAACGTG	GCTATTGGCC	TGGTCAACGT	GGTGTGGTGC	660
CTGGCCTGGT	GCCTGTGAA	CCAGCGGCCG	CTGCCTCACG	TGCGCAAGTG	CGTGGTGGTG	720
GTCTGCTGC	TGCAGGGCT	TCCTGCTC	GAGCTGCTTG	ACTTCCACC	GCTCTTCTGG	780
GTCTGGATG	CCCATGGCAT	CTGGCACATC	AGCACCACATCC	CTGTCCACGT	CCTCTTTTC	840
AGCTTCTGG	AAGATGACAG	CCTGTACCTG	CTGAAGGAAT	CAGAGGACAA	GTTCAAGCTG	900

GACTGAAGAC	CTTGGAGCGA	GTCTGCCCA	GTGGGGATCC	TGCCCCCGCC	CTGCTGGCCT	960
CCCTTCTCCC	CTCAACCCTT	GAGATGATT	TCTCTTTCA	ACTTCTTGAA	CTTGGACATG	1020
AAGGATGTGG	GCCCAGAAC	ATGTGGCCAG	CCCACCCCT	GTGGGCCCTC	ACCCAGCCTTG	1080
GAGTCTGTT	TAGGGAAGGC	CTCCCAGCAT	CTGGGACTCG	AGAGTGGGCA	GCCCCTCTAC	1140
CTCCCTGGAGC	TGAACCTGGG	TGGAACGTAG	TGTGCTCTTA	GCTCTACCGG	GAGGACAGCT	1200
GCCCTGTTCC	TCCCCATCAG	CCTCCTCCCC	ACATCCCCAG	CTGCCCTGGCT	GGTCCTGAA	1260
GCCCTCTGTC	TACCTGGGAG	ACCAAGGACC	ACAGGCCCTA	GGGATAACAGG	GGGTCCCCCTT	1320
CTGTTACAC	CCCCCACCC	CCTCCAGGAC	ACCAACTAGGT	GGTGTGAG	GCTTGTCTT	1380
TGCCAGCCA	AGGTTCACCG	CGATTCTCCC	CATGGGATCT	TGAGGGACCA	ACCTGCTGGG	1440
ATTGGGAAGG	AGTTTCAACC	TGACCRRTG	CCTAGCCAGG	TTCCCAAGGAG	GCCTCACCAT	1500
ACTCCCTTTC	AGGGCCAGGG	CTCCAGCAAG	CCAGGGCTA	GGATCCTGTC	CTGCTGCTG	1560
GTGAGAGCC	TGCCACCGT	TGTCGGGAGT	GTGGGCCAGG	CTGAGTGCAT	AGGTGACAGG	1620
GCCGTGAGCA	TGGGCTGGG	TGTGTGAG	CTCAGGCACT	AGGTGCGCAG	TGTGGAGACG	1680
GGTGTGTCG	GGGAAGAGGT	GTGGCTCAA	AGTGTGTTGT	GTGCAGGGGG	TKGGTGTGTT	1740
AAGCGTGGGT	TAGGGGAACG	TGTCGCGG	TGCTGGTGGG	CATGTGAGAT	GAGTGACTGC	1800
CGGTGAATGT	GTCCACAGGT	GAGAGGTTG	AGCAGGATGA	GGGAACTCTG	TCACCATCAA	1860
TAATCACTTG	TGGAGCGCA	CTTGGCCAA	GACGCCACCT	GGCGGACAG	CAGGAGCTCT	1920
CCATGGCCAG	GCTGCTGTG	TGATGTTCC	CTGCTGGTG	CCCCTTTGCC	CGCCTCCCTG	1980
AAACCTACA	GGGTCCCCAC	ACAAACAGTC	CCTCCAGAAG	CAGCCCTCG	GAGGCAAGAGG	2040
AAGGAAAATG	GGGATGGCTG	GGGATCTCTC	CATCCCTCTT	TTCTCCTTG	CTTCGATGG	2100
CTGGCCTTCC	CCTCCAAAAC	CTCCATCCC	CTGCTGCCAG	CCCCTTTGCC	ATAGCCTGAT	2160
TTGGGGAGG	AGGAAGGGGC	GATTTGAGG	AGAAGGGAG	AAAGCTTATG	GCTGGCTG	2220
GTTCCTTCCC	TTCCACAGAGG	GTCTTACTGT	TCCAGGGTGG	CCCCAGGGCA	GGCAGGGGCC	2280
ACACTATGCC	TGCGCCCTGG	AAAGGTGAC	CCCTGCCATT	TACCAAGCAGC	CCTGGCATGT	2340
TCCTGCCCTCA	CAGGAATAGA	ATGGAGGGAG	CTCCAGAAAC	TTTCCATCCC	AAAGGCAGTC	2400
TCCTGGTTG	AAGCAGACTG	GATTTTGCT	CTGCCCTGCA	CCCCCTGTCC	CTCTTTGAGG	2460
GAGGGGAGCT	ATGCTAGGAC	TCCAACTCA	GGGAGCTGGG	TGGCTGCGC	TAGCTTCTTT	2520
TGATACTGAA	AACCTTAAG	GTGGGAGGT	GGCAAGGGAT	GTGCTTAATA	AATCAATTCC	2580
AAGCCTCAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAA			

BFG7 Protein sequence (SEQ ID NO:17)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 117-133, 179-195, 211-227, 235-251, 266-282, 296-312; PFAM domains: none; Summary: A type III membrane protein of unknown function; is adjacent to HER2 on the genome, and its overexpression in breast cancer is highly correlated with HER2 expression; may be used to predict HER2 overexpression and amplification.

RVDPRVRGER	MAGLAARLVL	LAGAAALASG	SQGDREPVYR	DCVLQCEEQN	CSGGALNHFR	60
SRQPIYMSLA	GWTCRDDCKY	ECMWVTVGLY	LQEGHKVPQF	HGKWPFSRFL	FFQEPAVASA	120
SFLNGLASLV	MLCRYRTFVP	ASSPMYHTCV	AFAWVSLNAW	FWSTVFHTRD	TDLTEKMDYF	180
CASTVILHSI	YLCCVRTVGL	QHPAVVSAFR	ALLLLMLTVH	VSYLSLIRFD	YGYNLVANVA	240
IGLVNVWWL	AWCLWNQRRL	PHVRKCVVVV	LLLQLGLSLE	LLDFPPLFWV	LDAHAIWHIS	300
TIPVHVLFSS	FLEDDSLYLL	KESEDKFKLD				

BCN4 DNA sequence (SEQ ID NO:18)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Nucleic Acid Accession #: n/a; Coding sequence: 143-874 (start and stop codons underlined)

GGGAGGGAGA	GAGGCGCGCG	GGTGAAGAGC	GCATTGATGC	AGCCTGCGC	GGCCTCGGAG	60
CGCGCCGGAG	CCAGACGCTG	ACCACGTTCC	TCTCTCGGT	CTCCTCCGCC	TCCAGCTCCG	120
CGCTGCCCG	CAGCCGGAG	<u>CCAT</u> CGCACC	CCAGGGCCC	GCCGCTCTCC	CGCAGCGGCT	180
CCGCGGCCCTC	CTGCTGCTCC	TGCTGCTGCA	GCTGCCCGCG	CCGTCGAGCC	CCTCTGAGAT	240
CCCCAAGGGG	AAGCAAAGG	CGCAGCTCCG	GCAGAGGGAG	GTGGTGGACG	TGTATAATGG	300
AATGTGTTA	CAAGGGCCAG	CAGGAGTGC	TGGTCGAGAC	GGGAGCCCTG	GGGCCATATGG	360
CATTCCGGT	ACACCTGGG	TCCCAGGTG	GGATGGATT	AAAGGAGAAA	AGGGGAATG	420
TCTGAGGGAA	AGCTTTGAGG	AGTCTGGAC	ACCCAACTAC	AAGCAGTGT	CATGGAGTT	480
ATTGAATTAT	GGCATAGATC	TTGGGAAAT	TGCGGAGTGT	ACATTTACAA	AGATGCGTTC	540
AAATAGTGCT	CTAAGAGTT	TGTTCACTG	CTCACTTCGG	CTAAAATGCA	GAATGCATG	600
CTGTCAGCGT	TGGTATTTC	CATTCAATGG	AGCTGAATGT	TCAGGACCTC	TTCCCATGTA	660
AGCTATAATT	TATTTGGACC	AAGGAAGCCC	TGAAATGAAT	TCAACAAATT	ATATTGATCG	720
CACTTCTTCT	GTGGAAGGAC	TTTGTGAAGG	AATTGGTGCT	GGATTAGTGG	ATGTTGCTAT	780
CTGGGTTGGC	ACTTGTTCAG	ATTACCCAA	AGGAGATGCT	TCTACTGGAT	GGAAATCAGT	840
TTCTCGCATC	ATTATTGAG	AACTACCAA	AAATATGCTT	TAATTTGAT	TTGCTACCTC	900
TTTTTTATT	ATGCTTGGAA	ATGTTCACT	TAATAGACAT	TTTAAATAAG	TTTATGTATA	960
CATCTGAATG	AAAAGCAAAG	CTAAATATGT	TTACAGACCA	AAGTGTGATT	TCACACTGTT	1020
TTAAATCTA	GCATTATTCA	TTTGCTTC	ATCAAAAGTG	GTTTCAATAT	TTTTTTAGT	1080
TGGTTAGAAT	ACTTTCTCA	TAGTCACATT	CTCTCAACCT	ATAATTTGGA	ATATTGTTGT	1140
GGTCTTTGT	TTTTCTCTT	AGTATAGCAT	TTTAAAAAAA	ATATAAAAGC	TACCAATCTT	1200
TGTACAATT	GTAAATGTTA	AGAATTTTTT	TTATATCTGT	AAATAAAAAA	TTATTTCCAA	1260
CAACCTAAA	AAAAAAAAAA	AAAAAAA				

BCN4 Protein sequence (SEQ ID NO:19)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Protein Accession #: n/a; Predicted Signal sequence: 1-30; TM domains: none; PFAM domains: none; Summary: a secreted protein; has a mouse orthologue (see sequence below).

MRPQGPAASP QRLRGLLLLL LLQLPAPSSA SEIPKGKQKA QLRQREVVDL YNGMCLQGPA 60
GVPGRDGSPG ANGIPGTPGI PGRDGFKGEK GECLRESFEE SWTPNYKQCS WSSLNYGIDL 120
GKIAECTFTK MRSNSALRVL FSGSLRLKCR NACCQRWYFT FNGAECGPL PIEAIYLDQ 180
GSPEMNSTIN IHRRTSSVEGL CEGIGAGLVD VAIWVGTCSD YPKGDASTGW NSVSRIIEE 240
LPK

Mouse BCN4 Protein sequence (SEQ ID NO:20)
Gene name: ESTs; Unigene number: Mm.41556

XXXXAAPPQL LLGLFLVLLL LLQLSAPSSA SENPKVKQKA LIRQREVVDL YNGMCLQGPA 60
GVPGRDGSPG ANGIPGTPGI PCQDGFKGEK GECLRESFEE SWTPNYKQCS WSSLNYGIDL 120
GKIAECTFTK MRSNSALRVL FSGSLRLKCR NACCQRWYFT FNGAECGPL PIEAIXXXX 180
XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXSD YPKGDAYTGW DSVSRIIEE 240
LPK

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In the Specification:

Paragraph (Table 1) beginning at line 1 of page 94 has been amended as follows (see attached pages 94-103):

SF 1281429 v1

Table 1

BCA4 DNA sequence (SEQ ID NO:1)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset Accession #: D13666; Nucleic Acid Accession #: NM_006475; Coding sequence: 12-2522 (start and stop codons underlined)

AGAGACTCAA	GATGATTCCC	TTTTACCCA	TGTTTCTCT	ACTATTGCTG	CTTATTGTTA	60
ACCCATAAA	CGCCAACAA	CATTATGACA	AGATCTGGC	TCATAGTCGT	ATCAGGGTC	120
GGGACCAAGG	CCCAAATGTC	TGTGCCCTC	AACAGATTT	GGGCACCAAA	AAGAAATACT	180
TCAGCACTTG	TAAGAACTGG	TATAAAAAGT	CCATCTGTGG	ACAGAAAACG	ACTGTTTAT	240
ATGAATGTTG	CCCTGGTT	ATGAGAAATGG	AAGGAATGAA	AGGCTGCCCA	GCAGTTTGC	300
CCATTGACCA	TGTTTATGGC	ACTCTGGGC	TCTGGGGAGC	CACCACAAACG	CAGCCTATT	360
CTGACGCTC	AAAATGAGG	GAGGAGATCG	AGGGAAAGGG	ATCCTTCACT	TACTTTGCAC	420
CGAGTAATGA	GGCTTGGGAC	AACTTGGATT	CTGATATCCG	TAGAGGTTTG	GAGAGCAACG	480
TGAATGTTGA	ATTACTGAAT	GCTTACATA	GTCACATGAT	TAATAAGAGA	ATGTTGACCA	540
AGGACTTAA	AAATGGCATG	ATTATTCTT	CAATGTATAA	CAATTGGGG	CTTTCTATTA	600
ACCATTATCC	TAATGGGTT	GTCACTGTTA	ATTGTGCTCG	AATCATCCAT	GGGAACCCAGA	660
TTGCAACAAA	TGGTGTGTC	TACATGCTT	ACCGTGTGCT	TACACAAATT	GGTACCTCAA	720
TTCAAGACTT	CATTGAAGCA	GAAGATGACC	TTTCATCTTT	TAGAGCAGCT	GCCATCACAT	780
CGGACATATT	GGAGGCCCTT	GGAGAGACG	GTCACTTCAC	ACTCTTGTGCT	CCCACCAATG	840
AGGCTTTG	GAAACTTCA	CGAGGTGTC	TAGAAAGGTT	CATGGGAGAC	AAAGTGGCTT	900
CCGAAGCTCT	TATGAGTAC	CACATCTAA	ATACTCTCA	GTGTTCTGAG	TCTATTATGG	960
GAGGAGCAGT	CTTGAGACG	CTGGAAGGA	ATACAAATTG	GATAGGATGT	GACGGTGACA	1020
GTATAACAGT	AAATGGAATC	AAAATGGTA	ACAAAAAGGA	TATTGTGACA	AATAATGGTG	1080
TGATCCATT	GATTGATCAG	GTCCTAATT	CTGATTCTGC	CAAACAAGTT	ATTGAGCTGG	1140
CTGGAAAACA	GCAAACACC	TTCACGGATC	TTGTGGCCCA	ATTAGGCTTG	GCATCTGCTC	1200-
TGAGGCCAGA	TGGAGAATAC	ACTTTGCTGG	CACCTGTGAA	TAATGCATT	TCTGTATGATA	1260
CTCTCAGCAT	GGTTCAGCGC	CTCCTTAAAT	TAATTCTGCA	GAATCACATA	TTGAAAGTAA	1320
AAAGTGGCCT	TAATGAGCTT	TACAACGGG	AAATACTGGA	AACCATCGGA	GGAAACACAGC	1380
TCAGAGTCTT	CGTATATCGT	ACAGCTGTCT	GCATTGAAAA	TTCATGCATG	GAGAAAGGGA	1440
GTAAGCAAGG	GAGAACCGGT	GCGATTCA	TATTCGCGA	GATCATCAAG	CCAGCAGAGA	1500
AATCCCTCCA	TGAAAAGTTA	AAACAAGATA	AGCCTTTAG	CACCTTCC	AGCCTACTTG	1560
AAAGCTGCAGA	CTTGAAGAG	CTCTGACAC	AACCTGGAGA	CTGGACATTA	TTTGTGCCAA	1620
CCAATGATGC	TTTAAGGGG	ATGACTAGTG	AAGAAAAAGA	AATTCTGATA	CGGGACAAAA	1680
ATGCTCTTCA	AAACATCATT	CTTTATCACC	TGACACCAGG	AGTTTTCATT	GGAAAGGAT	1740
TTGAACCTGG	TGTTACTAAC	ATTTAAAGA	CCACACAAGG	AAGCAAAATC	TTTCTGAAAG	1800
AACTAAATGA	TACACTCTG	GTGAATGAAT	TGAATCATAA	AGAATCTGAC	ATCATGACAA	1860
CAATGGTGT	AATTCTATGTT	GTAGATAAA	TCTCTATCC	AGCAGACACA	CCTGTTGGAA	1920
ATGATCAACT	GCTGAAAT	CTTAAATAAT	TAATCAAAT	CATCCAAATT	AAGTTGTTTC	1980
GTGTTGACAC	CTTCAAAAGA	ATCCCCGTGA	CTGCTATAC	AACTAAAATT	ATAACCAAAG	2040
TTGTGGAACC	AAAAATTAAA	GTGATTGAAG	GCAGCTTCA	GCCTATTATC	AAAATGAAAG	2100
GACCCACACT	AAACAAAGTC	AAAATGAAG	GTGAACCTGA	ATTCAGACTG	ATTAAGAAG	2160
GTGAAACAAT	AACTGAAGT	ATCCATGGAG	AGCCAATTAT	TAAAAAATAC	ACCCAAATCA	2220
TTGATGGAGT	GCCTGTGAA	ATAACTGAA	AAGAGACAGC	AGAGAACCGA	ATCATTACAG	2280
GTCCTGAAAT	AAAATACACT	AGGATTCTA	CTGGAGGTGG	AGAAACAGAA	GAAACTCTGA	2340
AGAAATTTGT	ACAAGAAGAG	GTCACCAAGG	TCACCAAATT	CATTGAAGGT	GGTGATGGTC	2400
ATTTATTTGA	AGATGAAGAA	ATTTAAAGAC	TGCTTCAGGG	AGACACACCC	GTGAGGAAGT	2460
TGCAAGCCAA	AAAAAAAGTT	CAAGGTTCTA	GAAGACGATT	AAGGGAAGGT	CGTTCTCAGT	2520
GAAAATCCAA	AAACCAAGAA	AAAATGTTA	TACAACCCCTA	AGTCATAAAC	CTGACCTTAG	2580
AAAATGTGA	GAGCCAAGTT	GACTTCAGGA	ACTGAAACAT	CAGCACAAAG	AAGCAATCAT	2640
CAAATAATTC	TGAACACAAA	TTTAAATT	TTTTTCTGA	ATGAGAAACA	TGAGGAAAT	2700
TGTGGAGTTA	GCCTCTGTG	GTAAAGGAAT	TGAAGAAAAT	ATAACACCTT	ACACCCCTTT	2760
TCATCTTGCAC	ATTTAAAGTT	CTGGCTAACT	TTGGAATCCA	TTAGAGAAA	ATCCTTGTCA	2820
CCAGATTCTAT	TACAATTCAA	ATCGAAGAGT	TGTTGACTGT	TATCCCATG	AAAAGACCGA	2880
GCCTTGTATG	TATGTTATG	ATACATAAA	TGCAACGAAG	CCATTATCTC	TCCATGGAA	2940
GCTAAGTTAT	AAAAATAGGT	GCTTGTGTA	CAAAACTTTT	TATATCAAAA	GGCTTTGCAC	3000
ATTTCTATAT	GAGTGGGTTT	ACTGGTAAAT	TATGTTATTT	TTTACAACTA	ATTTGTA	3060
CTCAGAATGT	TTGTCATATG	CTTCTTGC	TGCAATATTTT	TTAATCTCAA	ACGTTCAAT	3120
AAAACCATT	TTCAGATATA	AAGAGAATTA	CTTCACAAATTG	AGTAATTCA	AAAAACTCAA	3180
GATTAAAGTT	AAAAAGTGGT	TTGGACTTGG	GAA			

BCA4 Protein sequence (SEQ ID NO:2)

Gene name: osteoblast specific factor 2 (periostin); Unigene number: Hs.136348; Probeset Accession #: D13666; Protein Accession #: NP_006466; Predicted Signal sequence: 1-21; TM domains: none; PFAM domains: fasciclin_domains: 94-232, 234-367, 496-630; Summary: a secreted protein involved in adhesion and osteoblast development; may participate in preferential metastasis of breast cancer to bone.

MIPFLPMFSL	LLLLIVNPIN	ANNHYDKILA	HSRIRGRDQG	PNVCALQQIL	GTKKKYFSTC	60
KNWYKKSICG	QKTTVLYECC	PGYMRMEGMK	GCPAVLPIDH	VYGTGLIVGA	TTTORYSDas	120
KLREEIEGKG	SFTYFAPSNE	AWDNLDSDIR	RGLESNVNE	LLNALHSHMI	NKRMLTKDLK	180
NGM1IPSMDN	NLGLFINHYP	NGVVTVN CAR	IIHGNQIATN	GVVHVIDRVL	TQIGTSIQDF	240
IEAEDDLSSF	RAAAITSDIL	EALGRDGHT	LFAPTN EAFe	KLPRGVLERF	MGDKVASEAL	300
MKYHILNTLQ	CSESIMGGAV	FETLEGNTIE	IGCDGSITV	NGIKMVKKD	IVTNNGVIHL	360
IDQVLIPDSA	KQVIELAGKQ	QTTFDLVAQ	I GLASALRPD	GEYTL LAPVN	NAFSDDTLSM	420

VQRLKLILQ NHILKVKVGL NELYNGQILE TIGGQLRHF VYRTAVCIEN SCMEKGSKQG	480
RNGAIHIFRE IIKPAEKSILH EKLQDKRFS TFLSLLAAD LKELLTQPGD WTLFVPTNDA	540
FKGMTSEEKE ILIRDKNALQ NIILYHLTPG VFIGKGFEPE VTNLKTTQG SKIFLKEVND	600
TLLVNELSK ESDIMTTNGV IHVVDKLLYP ADTPVGNDQL LEILNKLIKY IQIKFVRGST	660
FKEIPVTVYT TKIITKVVEP KIKVIEGSLQ PIKTEGPTL TKVKIEGEPE FRLIKEGETI	720
TEVIHGEPII KKYTKIIDGV PVEITEKTR EERIITGPEI KYTRISTGGG ETEETLKLL	780
QEVTKVTKF IEGGDGHLF DEEIKRLLQG DTPVRKLQAN KKVQGSRRRL REGRSQ	

BCA7 DNA sequence (SEQ ID NO:3)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset Accession #: Z29083; Nucleic Acid Accession #: NM_006670; Coding sequence: 85-1347 (start and stop codons underlined)

CCGGCTCGCG CCCTCCGGGC CCAGCCTCCC GAGCCTTCGG AGCGGGCGCC GTCCCAGCCC	60
AGCTCGGGGG AAACCGCGAGC CGCGATGCCT GGGGGGTGCT CCCGGGGCCC CGCCGCCGGG	120
GACGGGCGTC TGCGGCTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCCTCGTCT	180
TCTCCCACCT CCTCGGCATC CTCCCTCTCC TCCCTCGGCG CGTTCTGGC TTCCGCCGTG	240
TCCGCCAGC CCCCGCTGCC GGACCACTGC CCCCGCTGT GCGAGTGCTC CGAGGCAGCG	300
CGCACAGTCA AGTGCCTAA CCGCAATCTG ACCGAGGTGC CCAACGGACCT GCCCGCTTAC	360
GTGCGCAACC TCTTCTTAC CGGCAACCG CTGGCCGCG TCCCTGCCG CGCCTTCGCC	420
CGCCGGCCGC CGCTGGCGGA GCTGGCCGCG CTCAACCTCA CGGGCAGCGG CCTGGACGAG	480
GTGCGGGGGG GGGCGCTCGA GCATCTGCC AGCTCGGCC AGCTCGACCT CAGCCACAAC	540
CCACTGGCCG ACCTCAGTCC CTTCGCTTC TCAGGGCAGCA ATGCCAGCGT CTCGGCCCCC	600
AGTCCCCCTG TGGAACTGAT CCTGAACAC ATCGTGCCTCC CTGAAGATGA GCGGCAGAAC	660
CGGAGCTTCG AGGGCATGGT GGTGGCGGCC CTGCTGGCGG GCCCTGCACT GCAGGGGCTC	720
CGCCGCTTGG AGCTGGCCAG CAACCACTTC CTTACCTGC CGGGGGATGT GCTGGCCCAA	780
CTGCCAGGCC TCAGGCACCT GGACTTAAGT ATAATTCTGC TGTTGAGCCT GACCTACGTG	840
TCCTCCCGCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC	900
CTTCACAATG GCACCCCTGGC TGAGTTGCAA GGTCTACCCC ACATTAGGGT TTTCTGGAC	960
AACAATCCCT GGGTCTCGA CTGCCACATG GCAGACATGG TGACCTGCCT CAAGGAACA	1020
GAGGTAGTGC AGGGCAAAA CGGGCTCACC TGTCATATC CGGGAAAAAT GAGGAATCGG	1080
GTCTCTTGG AACTCAACAG TGCTGACCTG GACTGTGACC CGATTCTTC CCCATCCCTG	1140
CAAACCTCTT ATGTCCTTCG GGGTATTGTT TTAGCCCTGA TAGGCGTAT TTTCTCCTG	1200
GTTTGTATT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAACATCAG AGATGCCTGC	1260
AGGGATCACA TGGAAAGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA	1320
AACCTCAGTT CTAACTCGG TGTCAGAGG ATATTAGAGG ACAGACCAAG GACAACCTG	1380
CATGAGATGT AGACCTAAC TTTATCTCTA CTAGGCTTGC TCCACTTCA TCCTCCACTA	1440
TAGATACAAAC GGACTTTGAC AAAAGCAGT GAAGGGGATT TGCTTCCTG TTATGAAAG	1500
TTTCTCGGTG TGTTCTGTT ATGTAAGACG ATGAAACAGTT GTGTATAGT TTTTACCCCTC	1560
TTCTTTTCT TGGAACTCCT CAACACGTAT GGAGGGATT TTCAGGTTTC AGCATGAACA	1620
TGGGCTTCTT GCTGTCTGTC TCTCTCTCA TAGCTCTCAA GTGTAGCAA GTGTACCCAC	1680
ACAGATAGCA TTCAACAAA GCTGCTCAA CTTTTCGAG AAAAATACTT TATTCTAAA	1740
TATCAGTTT ATTCTCATGT ACCTAAGTT TGAGAAAAT AATTGCATCC TATAAAACTGC	1800
CTGCAGACGT TAGCAGGCTC TTCAAAATAA CTCCATGGTG CACAGGAGCA CCTGCATCCA	1860
AGAGCATGCT TACATTTAC TGTTCTGCAT ATTACAAAAA ATAACATTGCA ACTTCATAAC	1920
TTCTTTGACA AAGTAAATTA CTTTTTGAT TGCACTTTAT ATGAAAATGT ACTGATTTT	1980
TTTTAATAAA CTGCATCGAG ATCCAACCGA CTGAATTGTT AAAAAGAAAA AAAATAAAG	2040
ATTCTTAAAAA GAA	

BCA7 Protein sequence (SEQ ID NO:4)

Gene name: 5T4 oncofetal trophoblast glycoprotein; Unigene number: Hs.82128; Probeset Accession #: Z29083; Protein Accession #: NP_006661; Predicted Signal sequence: 1-32; Predicted TM domains: 357-373; PFAM domains: leucine rich repeats: 61-90, 119-142, 143-166, 235-258, 259-282, 294-345; Summary: a type 1a TM protein of unknown function, detected in multiple cancers, with highest expression in breast cancer.

MPGGCSRGPAGDGRLRLAR LALVLLGWVS SSSPTSSASS FSSSAPFLAS AVSAQQPLPD	60
QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRNLFLTG NQLAVLPAGA FARRPPLAEL	120
AALNLGSRL DEVRAFAEH LPSLRQLDLS HNPLADLSPF AFSGSNASVS APSPLVELIL	180
NHIVPPEDER QNRSFEGMVV AALLLAGRALQ GLRRLELASN HFLYLPDRVL AQLPSLRHLD	240
LSNNSLVSLT YVSFRNLTHL ESLHLEDNAL KVLLHNGTIAE LQQLPHIRVF LDNNNPWCDC	300
HMADMVTWLK ETEVVQGKDR LTCAYPEKMR NRVLLELNSA DLDCDPILPP SLQTSYVFLG	360
IVLALIGAIF LLVLYLNRK IKKWMHNIRD ACRDHMEGYH YRYEINADPR LTNLSSNSDV	

BCX5 DNA sequence (SEQ ID NO:5)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Nucleic Acid Accession #: AF160477; Coding sequence: 225-1757 (start and stop codons underlined)

GGGGAGCTCG GAGCTCCCGA TCACGGCTTC TTGGGGTAG CTACGGCTGG GTGTGTAGAA	60
CGGGGCCGGG GCTGGGCTG GGTCCCCTAG TGAGACCAA GTGCGAGAGG CAAGAACTCT	120
GCAGCTTCTCT GCTTCTGGG TCAGTTCTTT ATTCAAGTCT GCAGCCGCT CCCAGGGAGA	180
TCTCGTGGAA ACTTCAGAAA CGCTGGCAG TCTGCCTTTC AACCATGCC CTGTCCCTGG	240
GAGCCGAGAT GTGGGGCCCT GAGCCCTGGC TGCTGCTGCT GCTACTGCTG GCATCATTTA	300
CAGGCGGTG CCCCGGGGT GAGCTGGAGA CCTCAGACGT GGTAACCTGT GTGCTGGGCC	360
AGGACCAAAC ACTGCCCTGC TTCTACCGAG GGGACTCCGG CGAGCAAGTG GGGCAAGTGG	420

CATGGGCTCG	GGTGGACGCG	GGCGAAGGCG	CCCAGGAACT	ACCGCTACTG	CACTCCAAAT	480
ACGGGCTTC	TGTGAGCCCCG	GCTTACGAGG	GCCGCGTGG	GCAGCCGCCG	CCCCCAGC	540
ACCCCTGGA	CGGCTCAGTG	CTCCTGCGCA	ACGCAGTGCA	GGCGGATGAG	GGCGAGTAGC	600
AGTGCCTGGT	CAGCACCTTC	CCCAGGGCA	GCTTCCAGGC	GGCGCTGCG	CTCCGAGTGA	660
TGGTGCCTCC	CTCTGCA	GTCAGGAGCA	GGCCAGCCCC	CAGCGTGACC	TGGGACACGG	720
CCCTGGCAGC	CTACAGTCCC	AGCCGTTCC	TCAAGCACTC	CCGCTCTGCT	GGCGTCACCT	780
AGGTCAAAGG	CACAACTCC	AGCCGTTCC	TCAAGCACTC	CCGCTCTGCT	GGCGTCACCT	840
CAGAGTTCCA	CTTGGTGCCT	AGCCGCAGCA	TGAATGGGCA	GCCACTGACT	TGTGTGGTGT	900
CCCATCCTGG	CCTGCTCCAG	GACCAAAGGA	TCACCCACAT	CCTCCACGCG	TCCTTCCTTG	960
CTGAGGCCTC	TGTGAGGGC	CTTGAAGAAC	AAAATCTGTG	GCACATTGGC	AGAGAAGGAG	1020
CTATGCTCAA	GTGCGTGTAG	GAAGGGCAGC	CCCCCTCC	ATACAACCTGG	ACACGGCTGG	1080
ATGGGCTCT	GCCCAGTGG	GTACGAGTGG	ATGGGACAC	TTTGGGCTTT	CCCCCACTGA	1140
CCACTGAGCA	CAGCGGCATC	TACGTCGCC	ATGTCAGCAA	TGAGTTCTCC	TCAAGGGATT	1200
CTCAGGTAC	TGTGGATGTT	CTTGACCCCC	AGGAAGACTC	TGGGAAGCAG	GTGGACCTAG	1260
TGTAGCCTC	GGTGGTGTG	GTGGGGTGTGA	TCGGCCACT	CTTGTCTCTG	CTTCTGGTGG	1320
TGGTGTGGT	GTCATGTC	CGATACCATC	GGCGCAAGGC	CCAGCAGATG	ACCCAGAAAT	1380
ATGAGGAGGA	GCTGACCC	ACCAGGAGA	ACTCCATCCG	GAGGCTGCAT	TCCCACATCACA	1440
CGGACCCAG	GAGCCAGCCC	GAGGAGAGTG	TAGGGCTGAG	AGCCGAGGGC	CACCCCTGATA	1500
GTCTCAAGGA	CAACAGTAGC	TGCTCTGTGA	TGAGTGAAGA	GCCCCGAGGGC	CGCAGTTACT	1560
CCACGCTGAC	CACGGTGAGG	GAGATAGAAA	CACAGACTGA	ACTGCTGTCT	CCAGGCTCTG	1620
GGCGGGCCGA	GGAGGAGGAA	GATCAGGATG	AAGGCATCAA	ACAGGGCATAG	AACCATTG	1680
TTCAGGAGAA	TGGGACCTCA	CCACGGCAAGC	TGGCATCTAC	ATCAATGGC	1740	
GGGGACACCC	GGTCTGACCC	AGGCCTGCT	CCCTCCCTA	GGCCTGGCTC	CTTCTGTGA	1800
CATGGGAGAT	TTTAGCTCAT	CTTGGGGGCC	TCCTTAACAA	CCCCCATTTC	TTGCGGAAGA	1860
TGCTCCCAT	CCCACGACT	GCTTGACCTT	TACCTCCAC	CCTTCTGTG	ATCGGGAGGG	1920
CTCCACCAAT	TGAGTCTCTC	CCACCATGCA	TGCAAGGTCAC	TGTGTGTG	CATGTGTGCC	1980
TGTTGACTG	TTGACTGACT	GTGTGTG	GGAGGGTGA	CTGTCCTG	AGGGGTGACT	2040
GTGTCCTGG	TGTGTATTAT	GCTGTCATAT	CAGAGTCAG	TGAACGTG	TGTATGTGCC	2100
ACGGGATTG	AGTGGTTGCG	TGGGCAACAC	TGTCAGGGTT	TGGCGTGTG	GTCATGTGCC	2160
TGTGTGTGAC	CTCTGCTGA	AAAAGCAGGT	ATTTTCTCG	ACCCCAAGGC	AGTATTAAATG	2220
ATGCAGAGGT	TGGAGGAGAG	AGGTTGAGAC	TGGGCTGAC	ACCCAGGTG	GCAGGCATAG	2280
CTGGAGCTGG	AATCTGCC	CGGTGTGAGG	GAACCTGCT	CTTACCACT	CGGAGCCATG	2340
GGGGCAAGTG	TGAAGCAGCC	AGTCCCTGG	TCAGCCAGAG	GCTTGAAC	TTACAGAAC	2400
CCTCTGCC	CTGGTGGCCT	CTGGGCTGC	TGCACTGACA	TATTTTCTGT	AAATATACAT	2460
GGCCCGGGAG	CTTCTTGCA	GAATACTGCT	CCGAATCACT	TTAATTTTT	TTCTTTTTT	2520
TTCTTGCCC	TTTCATTAG	TTGTATTTTT	TATTTTCTGT	TATTTTCTGT	TTTTTTAGA	2580
GATGGAGTGT	CACTATGTTG	CTCGAGCTGG	CCTTGAAC	CTGGGCTCAA	GCAATCTCC	2640
TGCCTCAGCC	TCCCCTAGTAG	CTGGGACTT	AAGTGTACAC	CACTGTGCC	GCTTGAATC	2700
CTTACGAAG	AGAAAAAAA	AATTAAAGAA	AGCCTTAA	TTTATCCAAT	GTAACTACT	2760
GGGATTGCTT	AAAGTGAGGC	CCCTCCAACA	CCAGGGGGTT	AATTCTGTG	ATTGTGAAG	2820
GGGCTACTTC	CAAGGCATCT	TCATGCAGGC	AGCCCTTGG	GAGGGCACCT	GAGAGCTGG	2880
AGAGTCTGAA	ATTAGGGATG	TGAGCTCTGT	GGTACTGAG	TAAGGTAAA	TTGCATCCAC	2940
CATTGTTGT	GATACTT	GGAAATTGCTT	GGACAGTGG	ACAAGGGCTC	CTGTTCAATA	3000
GTGGTGTG	GGAGAGAGG	AGCAGTGTG	ATAGACCGAG	AGAGTAGGAG	TTGAGGTGAG	3060
GTAAAGGAGG	TGCTGGGGT	GAGAAATGTC	CCTTCCCCC	TGGGTTTGG	ATCACTAATT	3120
CAAGGCTCTT	CTGGATGTTT	CTCTGGGTTG	GGGCTGGAGT	TCAATGAGGT	TTATTTTTAG	3180
CTGGCCACC	CAGATACACT	CAGCCAGAAT	ACCTAGATTT	AGTACCCAA	CTCTCTTAG	3240
TCTGAAATCT	GCTGGATTTC	TGGCTTAAGG	GAGAGGCTCC	CATCTTCGT	TCCCCAGCCA	3300
GCCTAGGACT	TCGAATGTGG	AGCCTGAAGA	TCTAAGATCC	AAACATGTAC	ATTTTATGTA	3360
AATATGTGCA	TATTTGTA	AAAATGATA	TTCTGTTTT	AAATAAACAG	ACAAAAC	3420
TTCAAAAAAA	AAAAAAAAA	AAAAAA	AAAAAA			

BCX5 Protein sequence (SEQ ID NO:6)

Gene name: LNIR; Unigene number: Hs.61460; Probeset Accession #: AA028028; Protein Accession #: AF160477; Predicted Signal sequence: 1-26; Predicted TM domains: 355-371; PFAM domains: IgSF_domain: 45-129, 162-225, 263-317; Summary: A type Ia TM protein; is a member of the immunoglobulin superfamily.

MPLSLGAEMW	GPEAWL	LLLASFTGRCP	AGELETSDVV	TVVLGQDAKL	PCFYRGDSGE	60
QVGQVAWARV	DAGEGAQELA	LLHSKYGLHV	SPAYEGRVEQ	PPPPRNPLDG	SVLLRNAVQA	120
DEGEYECRVS	TFPAGSFQAR	LRLRVMVPP	PSLNPGPALE	EGQGLTLAAS	CTAEGSPAPS	180
VTWDTEVKGT	TSSRSFKHSR	SAAVTSFHL	VPSRSRNMNQP	LTCVVSHPGL	LQDQRITHIL	240
HVSFLAEASV	RGLEDQNWLH	IGREGAMLKC	LSEGQPPPSSY	NWTRLDGPLP	SGVRVDGDTL	300
GFPPLTTEHS	GIYVCHVSNE	FSSRDLSQVT	DVLDPQEDSG	KQVDLVSASV	VVVGVIALL	360
FCLVVVVV	MSRYHRRKAQ	QMTQKYEEL	TLTRENSIRR	LHSHHTDPRS	QPEESVGLRA	420
EGHPDSDLKD	SSCSVMSEEP	EGRSYSTLTT	VREIETQTEL	LSPGSGRAEE	EEDQDEGIKQ	480
AMNHFVQENG	TLRAKPTGNG	IYINGRGLHV				

mouse BCX5 Protein sequence (SEQ ID NO:7)

Gene name: mouse_LNIR; Unigene number: n/a; Probeset Accession #: BF168327; Protein Accession #: n/a; Predicted Signal sequence: 1-27; Predicted TM domains: 346-362; PFAM domains: IgSF_domains:44-126, 166-221, 259-313; Summary: This is the mouse orthologue of human BCX5; it is a type Ia TM protein of unknown function.

MPLSLGAEMW	GPEAWL	LLLASFTGQYSA	GELETSDVVT	VVLGQDAKLP	CFYRGDPDEQ	60
VGQVAWARVD	PNEXYPGAGL	LHSKYGLHV	PAYEDRVEQX	XHETFRRSVL	LRNAVQAEG	120

EYECRVSTFP SGSFQARMRL RVLVPLPSL NPGPPLLEEGQ ADVAASCTAE GSPAPSVTWD 180
 TEVKGTQSSR SFTHPRSAV TSEFHLPVSR SMNGQPLTCV VSHPGQLQDR RITHTLQVAF 240
 LAEASVRGLE DQNLWQVGRE GATLKCLSEG QPPPQKYNWTR LDGPLPSGVR VKGDTLGFPP 300
 LTTEHSGVYX CHVSNELSSR DSQVTVEVLD PEDPGKQVDL VSASVIIIVGV IAALLFCLLV 360
 VVVVLMRSYH RRKAQQMTQK YEEELTLTRE NSIRRLHSHH SDPRSQPEES VGLRAEGHPD 420
 SKDNSSCSV MSEEPEGRSY STLTTVREIE QTTELLSPGS GRTEEDDDQD EGIKQAMNHL 480
 CRKMGP

BCZ6 DNA sequence (SEQ ID NO:8)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Nucleic Acid Accession #: NM_002184;
 Coding sequence: 256-3012 (start and stop codons underlined)

GAGCAGCCAA	AAGGCCCGCG	GAGTCGCGCT	GGGCCGCCCC	GGCGCAGCTG	AACCGGGGGC	60
CGCGCCTGCC	AGGCCGACGG	GTCGCCCCA	GCCTGGCGCC	AAGGGGTTCG	TGCGCTGTGG	120
AGACGCGGAG	GGTCGAGGGCG	GCGCGGCCG	AGTGAACCCC	AATGGAAAAA	GCATGACATT	180
TAGAAGTAGA	AGACTTAGCT	TCAAATCCCT	ACTCCTTCAC	TTACTAATT	TGTGATTG	240
AAATATCCGC	<u>GCAAGATGTT</u>	GACGTTGCAG	ACTTGGGTAG	TGCAAGCCTT	GTTTATTTTC	300
CTCACCACTG	AATCTACAGG	TGAACCTCTA	GATCCATGTG	GTTATATCA	TCCCTGAATCT	360
CCAGTTGTAC	AACCTCATTC	TAATTTCACT	GCAGTTGTG	TGCTAAAGGA	AAAATGTATG	420
GATTATTTTC	ATGTAATG	TAATTCAATT	GTCTGGAAA	CAAACATT	TACTATTCT	480
AAGGAGCAAT	ATACATCAT	AAACAGAAC	GCATCCAGTG	TCACCTTAC	AGATATAGCT	540
TCAATTAAATA	TTCAAGCTCAC	TTGCAACATT	CTTACATTG	GACAGCTTGA	ACAGAATGTT	600
TATGGAATCA	CAATAATTTC	AGGCTTGCCT	CCAGAAAAAC	CTAAAATT	GAGTTGCATT	660
GTGAACGAGG	GGAAAGAAAAT	GAGGTGAG	TGGGATGGTG	GAAGGGAAAC	ACACTTGGAG	720
ACAAACTCTA	CTTTAAATC	TGAATGGGC	ACACACAAGT	TTGCTGATTG	CAAAGCAAAA	780
CGTGACACCC	CCACCTCATG	CACTGTTGAT	TATTCTACT	TGTTTTG	CAACATTGA	840
GTCTGGGTAG	AAGCAGAGAA	TGCCCCGGG	AAGGTTACAT	CAGATCATAT	CAATTTGAT	900
CCTGTATATA	AACTGAAGCC	CAATCCGCCA	CATAATTAT	CAGTGTCA	CTCAGAGGAA	960
CTGTCTAGTA	TCTTAAATT	GACATGGACC	AACCCAAGTA	TTAAGACTGT	TATAATACTA	1020
AAATATAACA	TTCAATATAG	GACCAAAAGT	GCCTCAACTT	GGAGCCAGAT	TCCTCCTGAA	1080
GACACAGCAT	CCACCCGATC	TTCATTCACT	GTCCAAAGACC	TTAACCTT	TACAGAATAT	1140
GTGTTTAGGA	TTCGCTGT	GAAGGAAGAT	GGTAAGGGAT	ACTGGAGTGA	CTGGAGTGA	1200
GAAGCAAGTG	GGATCACTTA	TGAAGATAGA	CCATCTAAAG	CACCAAGTT	CTGGTATAAA	1260
ATAGATCCAT	CCCATACTCA	AGGCTACAGA	ACTGTACAAAC	TCGTGTGAA	GACATTGCCT	1320
CCTTTGAAAG	CCAATGGAAA	AATCTTGGAT	TATGAAAGTGA	CTCTCACAAG	ATGGAAATCA	1380
CATTACAAAC	ATTACACAGT	TAATGCCACA	AAACTGACAG	TAAATCTC	AAATGATCGC	1440
TATCTAGCAA	CCCTAACAGT	AAGAAATCTT	GTGCGAAAT	CAGATGCAGC	TGTTTTAACT	1500
ATCCCTGCCT	GTGACTTTCA	AGCTACTTCAC	CCTGTAATGG	ATCTTAAAGC	ATTCCCCAAA	1560
GATAACATGC	TTGGGGTGG	ATGGACTACT	CCAAGGGAAAT	CTGTAAGAA	ATATATAACTT	1620
GAGTGGTGTG	TGTTATCAGA	AAAGGCCACC	TGATCACAG	ACTGGCAACA	AGAAGATGGT	1680
ACCGTGCATC	GCACCTATT	AAGAGGAAC	TTAGCAGAGA	GCAAATGCTA	TTTGATAACA	1740
GTACTCCAG	TATATGCTGA	TGGACCAAGA	AGCCCTGAAT	CCATAAAGGC	ATACCTTAAA	1800
CAAGCTCCAC	CTTCCAAAGT	ACTACTGTG	CGGACAAAAA	AAGTAGGGAA	AAACGAAGCT	1860
GTCTTAGAGT	GGGACCAACT	TCCTGTTGAT	GTTCAGAAATG	GATTTATCA	AAATTATACT	1920
ATATTTTATA	GAACCATCAT	TGGAATGAA	ACTGCTGTGA	ATGTGGATT	TTCCCACACA	1980
GAATATACAT	TGTCCTCTT	GAAGTGTGAC	ACATITGTACA	TGGTACGAAT	GGCAGCATAC	2040
ACAGATGAAG	GTGGGAAGAG	TGGTCAGAA	TTCACTTTTA	CTACCCCCAA	GTTGCTCAA	2100
GGAGAAATTG	AAGGCATAGT	CGTGCCTGTT	TGCTTAGCAT	TCCTATTGAC	AACTCTCTG	2160
GGAGTGTGT	TCTGTTAA	TAAGCGAGAC	CTAATTAAAA	AACACATCTG	GCCTAATGTT	2220
CCAGATCCTT	CAAAGAGTC	TATTGCCAG	TGGTCACCTC	ACACTCCCTC	AAGGCACAAT	2280
TTTAATTCAA	AAGATCAAAT	GTATTCACTG	GGCAATTTC	CTGATGTAAG	TGTTGTGGAA	2340
ATAGAAGCAA	ATGACAAAAA	GCCTTTCCA	GAAGATCTGA	AATCATTGGA	CCTGTTCAA	2400
AAGGAAAAATG	TTAATACTGA	AGGACACAGC	AGTGGTATTG	GGGGGTCTTC	ATGCATGTCA	2460
TCTTCTAGGC	CAAGCATTC	TAGCACTGAT	GAAAATGAAT	CTTCACAAA	CACTTCGAGC	2520
ACTGTCCAGT	ATTCTACCGT	GGTACACAGT	GGCTACAGAC	ACCAAGTTC	GTCACTCCAA	2580
GTCTTCTCAA	GATCCGAGTC	TACCCAGCCC	TTGTTAGATT	CAGAGGAGCG	GCCAGAAAGAT	2640
CTACAAATTAG	TAGATCTGAT	AGATGGCGAT	GATGGTTATT	TGCCCCAGGA	ACAGTACTTC	2700
AAACAGAAACT	GCAGTCAGCA	TGAATCCAGT	CCAGATATT	CACATTGTA	AAGCTCAAAG	2760
CAAGTTTCAT	CACTCAATG	GGAAGATT	TTAGACTTA	AACAGCAGAT	TTCAGATCAT	2820
ATTCACAAAT	CCTGTGGATC	TGGGAAATG	AAAATGTTTC	AGGAAGTT	TGCAAGCAGAT	2880
GCTTTGGTC	CAGGTACTGA	GGGACAAAGTA	GAAAGATTG	AAACAGTTG	CATGGAGGCT	2940
CGGACTGTAG	AAGGCATGCC	TAAGGATTAC	TTACCAACAGA	CTGTACGGCA	AGGC GGCTAC	3000
ATGCCTCAGT	GAAGGACTAG	TAGTTCTG	TACAACCTCA	GCAGTACCTA	TAAAGTAAAG	3060
CTAAAATGAT	TTTATCTG	AATT				

BCZ6 Protein sequence (SEQ ID NO:9)

Gene name: IL-6 receptor beta chain (gp130; oncostatin M receptor); Unigene number:
 Hs.82065; Probeset Accession #: M57230 / AA406546; Protein Accession #: NP_002175; Predicted
 Signal sequence: 1-22; Predicted TM domains: 625-641; PFAM domains:
 fibronectin_type_III_domains: 222-311, 424-509, 519-606; Summary: A type I TM protein; it
 homodimerizes or heterodimerized to make a functional receptor for IL-6, oncostatin-M, IL-11,
 LIF, and CNTF.

MLTLQTWVHQ	ALFIFLTTES	TGEELLDPCGY	ISPESPVQQL	HSDNFTAVCVL	KEKCMDYFHV	60
NANYIVWKTN	HFTIPKEQYT	IINRTASSVT	FTDIASLNIQ	LTCNILTFQ	LEQNVYGITI	120

ISGLPPEKPK	NLSCIVNEGK	KMRCEWDGGR	ETHLETNFTL	KSEWATHKFA	DCKAKRDTPT	180
SCTVDYSTVY	FVNIEVWVEA	ENALGKVTS	HINFDPVYKV	KPNPPHNLSV	INSEELSSIL	240
KLTWTNPSIK	SVIILKYNIQ	YRTKDASTWS	QIPPEDTAST	RSSFTVQDLK	PFTEYVFIR	300
CMKEDGKGW	SDWSEEASGI	TYEDRPSKAP	SFWYKIDPSH	TQGYRTVQLV	WTKLPPFEAN	360
GKILDYEVTL	TRWKSHLQNY	TVNATKLTVN	LTNDRYLATL	TVRNLVGKSD	AAVLTIACD	420
FQATHPVMDL	KAFPKDNMLW	VEWTPPRESV	KKYIWTWCWL	SDKAPCITDW	QQEDGTVHRT	480
YLRGNLAEVK	CYLITVTPVY	ADGPGSPESI	KAYLKQAPPS	KGPTVRTKKV	GKNEAVLEWD	540
QLPVDVONGF	IRNYTIFYRT	IIGNETAVNV	DSSHTEYTLS	SLTSDTLYMV	RMAAYTDDEGG	600
KDGPEFTFTT	PKFAQGEIEA	IVVPVCL AFL	LTTLLGVLF	FNKRDLIKHH	IWPNVPDPSK	660
SHIAQWSPT	PPRHNFNSKD	QMYSDFGNFTD	VSVVEIEAND	KKPFPEDLKS	LDLFFKEKIN	720
TEGHSSGIGG	SSCMSSSRPT	ISSLDENESS	QNTSSTVQYS	TVVHSGYRHQ	VPSVQVFSRS	780
ESTQPILLDSE	ERPEDLQLVD	HVDGGDILP	ROQYFKQNSC	QHESSPDISH	FERSKQVSSV	840
NEEDFVRLKQ	QISDHISQSC	GSGQMKMFQE	VSAADAFGPG	TEGQVERFET	VGMEAATDEG	900
MPKSYLPQTV	RQGGYMPQ					

BFG4 DNA sequence (SEQ ID NO:10)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
 Nucleic Acid Accession #: AB020689; Coding sequence: 108-2777 (start and stop codons
underlined)

GAACTTATGT	AGCCTCATTA	TCCCGCTCCG	TGAGGTGACA	ATTGTGGAAA	AGGCAGACAG	60
CTCCAGTGTC	CTCCCCAGTC	CCTTATCAC	TCAGCACCCG	<u>AAACAGGATG</u>	ACCTTCCAT	120
TTGCCAACTT	GAAAGATAGA	GACTTTCTAG	TGCAAGAGGAT	CTCAGATTTC	CTGCAACAGA	180
CTACTTCAA	AATATATTCT	GACAAGGAGT	TTGCAGGAAG	TTACAACAGT	TCAGATGATG	240
AGGTGTACTC	TCGACCCAGC	AGCCTCGTCT	CCTCCAGGCC	CCAGAGAACG	ACGAGCTCTG	300
ATGCTGATGG	AGAGCGCAG	TTTAACCTA	ATGCAACAG	CGTCCCCAC	GCCACACAGA	360
CCCTGATGAC	CATGATATCGG	CGGGCGCTC	CCGAGGAGTT	CAACCCGAAA	TTGGCCAAG	420
AGTCTCTGAA	AGAGCAAGC	TGGAAGATT	ACTTTGCTGA	CTATGGGCAA	GGGATCTGCA	480
TGTACCGCAC	AGAGAAAACG	CGGGAGCTGG	TGTTGAAGGG	CATCCCGAG	AGCATGCGT	540
GGGAGCTCTG	GCTGCTGCTG	TCAGGTGCCA	TCAATGAGAA	GGCCACACAT	CCTGGGTACT	600
ATGAAGACCT	AGTGGAGAAG	TCCATGGGG	AGTATAATCT	CGCCACGGAG	GAGATTGAGA	660
GGGATTTACA	CCGCTCCCTT	CCAGAACACC	CAGCTTTCTA	GAATGAAATG	GGCATTGCTG	720
CACTAAGGAG	AGTCTTAAAC	GCTTATGCTT	TTGCAAATCC	CAACATAGGG	TATTGCCAGG	780
CCATGAATAT	TGTCACTTCA	GTGCTGCTGC	TTTATGCCAA	AGAGGAGGAA	GCTTTCTGGC	840
TGCTTGTGGC	TTTGTGTGAG	CGCATGCTCC	CAGATTACTA	CAACACAGA	GTTGTGGGTG	900
CACTGGTGG	CCAAGGTGTC	TTTGAGGAGC	TAGCACGAGA	CTACGCTCCC	CAGCTGTACG	960
ACTGCATGCA	AGACCTGGC	GTGATTTCA	CCATCTCCCT	GTCTTGTTC	CTCACACTAT	1020
TTCTCAGTGT	GATGCTT	GAGAGTCG	TTGTTGGTTG	TGACTGTTTC	TTCTATGAAG	1080
GAATTAAAGT	GATATTCCAG	TTGGCCCTAG	CTGTGCTGGA	TGCAAATGTG	GACAAACTGT	1140
TGAAC TGCAA	GGATGATGGG	GAGGCCATGA	CCGTTTTGGG	AAGGTATTTA	GACAGTGTGA	1200
CCAATAAAAGA	CAGCACACTG	CCTCCCATTC	CTCACCTCCA	CTCCTGCTC	AGCGATGATG	1260
TGGAACCTTA	CCCTGAGCTA	GACATTTTA	GACTCATCAG	AACCTCTCA	GAGAAATTG	1320
GAACATATCCG	GGCAGATTAG	ATTGAACAGA	TGAGATTCAA	ACAGAGACTG	AAAGTGTATCC	1380
AGACGCTGGA	GGATACTAAC	AAACGCAACG	TGGTACGAAC	CATTGTGACA	GAAACTTCCT	1440
TTACCATTTGA	TGAGCTGGAA	GAACATTATG	CTCTTTCAA	GGCAGAACAT	CTCACAGCT	1500
GCTACTGGGG	CGGGAGCAGC	AACCGCCTGG	ACCGGCATGA	CCCCAGCTG	CCCTACCTGG	1560
AACAGTATCG	CATTGACTTC	GAGCAGTTCA	AGGAATGTT	TGCTCTTCTC	TTTCTTGGG	1620
CATGTGGAAC	TCACTCTGAC	GTTCTGGCT	CCCGCTTGT	CCAGTTATTA	GATGAAAATG	1680
GAGACTCTT	GATTAACCT	CGGGAGTTG	TCTCTGGCT	AAGTGTGCA	TGCCATGGG	1740
ACCTCACAGA	GAAGCTCAA	CTCCTGTACA	AAATGCACGT	CTTGCCTGAG	CCATCCTCTG	1800
ATCAAGATGA	ACCA GATTCT	GCTTTGAGA	CAACTCAGTA	CTTCTTTGAA	GATATTACCC	1860
CAGAA GTAC	ACATGTTGTT	GGATTGGATA	GCAGAACAA	ACAGGGTGC	GATGATGGCT	1920
TTGTTACCGT	GAGCCTAAAG	CCAGAACAAAG	GGAAAGAGGC	AAATTCCCAA	GAAAATCGTA	1980
ATTATTTGAG	ACTGTGGACT	CGAACAAAATA	AAATCTAAC	AAAAGATGCA	AAGGATTTAC	2040
CCAAATTTAA	TCAGGGGAG	TTCATGAA	TGTTGAAGAC	AATGTATAAC	ATGTTCAGCG	2100
AAGACCCCAA	TGAGCAGGAG	CTGTACCATG	CCACGGCAGC	AGTGACCAGC	CTCCTGCTGG	2160
AGATTGGGG	GGTCGGCAAG	TTGTTCTGTT	CCCA CGCTGC	AAAGGAGGGC	GGGAGCGGAG	2220
GCAGTGGGCC	GTCCTGCAC	CAGGGCATCC	CAGCGCTGCT	CTTCCCGAAG	AAAGGGCCAG	2280
GCCAGCCTTA	CGTGGTGGAG	TCTGGTGGAG	CCCTGCCGCG	CAGCCTGGCC	CCCGACAGCG	2340
AGGAACACTC	CCTTGGAGGA	CAATGGAGG	ACATCAAGCT	GGAGGACTCC	TCGCCCCGGG	2400
ACAACGGGG	CTGCTCTCC	ATGCTGATCT	CTGACGACGA	CACCAAGGAC	GACAGCTCCA	2460
TGTCCTCATA	CTCGGTGCTG	AGTGGCGCT	CCACAGGAGG	GGACAAGCTG	CACTGGGAGG	2520
AAATCGGAA	GGACACGTC	CTGGTGGCGA	GGGGCCAGGG	CACGGCGGCA	CTGCCCCGG	2580
GCACCGCCT	GGACCGGGAC	GGGGCCATCA	CCTTCGAGCA	GTTCTGGCC	TCCCTCTTAA	2640
CTGAGCCTGC	CCTGGTCAAG	TACTTTGACA	AGCCCCGTGTG	CATGATGGCC	AGGATTACCA	2700
GTGCAAAAAA	CATCCGGATG	ATGGGCAAGC	CCCTCACCTC	GGCCAGTGAC	TATGAAATCT	2760
CGGCCATGTC	CGGCTGACAC	GGGCGCTTC	CCGGGGGAGT	GGGAGGAGAG	GGAGGGGAGG	2820
GATTTTTTAT	GTTCCTCTGT	GTTGAGTTT	TTCTTTCTTT	CTTTTAAAT	AAATATTAT	2880
TAGTACCTGG	AATTGAAGCC	TAGTGTGTT	ATAATGTAAT	TCAATGAAA	CTGTTGGAGA	2940
AAATTTAA	CACCTCAATG	TAGGTACATT	ACACTCTTGT	TGGGGGAGG	GGATTACCA	3000
GAATACAGTT	TATTTCTGTA	ATTCTAAAAA	ACAAAAAGAT	GAATCTGCA	GTGATATGTG	3060
TGTATTATAA	CTTATTAAATC	TTGCTGTTGA	GCTGTATACA	TGGTTAAA	AATAGTACTG	3120
TTAATGCTA	AGTAAGGCAG	CAGTCATTG	TGTTATTCA	CTTTTTAAAT	AAAATTAGAG	3180
CTGTAAGGAA	AATGAAAAGC	CACAAATGCA	AGACTGTTCT	TAATGGAAG	GCATAGTCAG	3240
CGAGGGTAA	TCCTATACCA	CTT TAGGAAG	TATTTAAAT	ATTTTTAAGA	TTTGAATAT	3300
ATTTCATAGA	AGTCCTCAT	TCAAAATCAT	ATTCCACAGA	TGTTCCCTT	CAAAGGGAAA	3360

ACATTTGGGG	TTCTAAACAG	TTATGAAAGT	AAGTGATTT	TACATGATT	CAGAATAACA	3420
CTTGATTGA	CCAATTTAGA	CAGATACCA	ACCAATTTC	CATTAAAGAA	ATTGTTCTGA	3480
TTATTTACGT	CAACTCATTA	GAATTCACTG	AAAAGTAACA	GTCCTTTGTC	ACAGAGAAC	3540
TGAAAGTAGC	AGCAAAGACA	GAGGGCTCAT	GACAGGTTT	TGCTTTTGCT	TTGCTTTG	3600
TTTTGAAAGA	GTAAAAGTAC	TGATGCTTCT	GATACTGGAT	GTTTAGCTTC	TTACTGC	3660
AACATAAGT	AAACAGTC	CTTTACATT	CCCGTATTCT	CCATAGATTG	AAGAAATT	3720
TACCACATAT	CGCATATGAC	CATCTTCCA	TCAAATCAAT	GTAGAGATAA	TGTAAACTGA	3780
AAAAAAATCT	GCAAGATAAT	GTAACGTGAAT	GT	TTTAAAAA	CAGAACTTGT	3840
AAAAGAATAG	TATGCTCTAT	TTCCTGAATG	GATGTCGAA	TGAAAGCTAG	CGCACCTGCA	3900
CTTTGAATT	TTGCTCTT	TTTATTACTG	TTATGATT	GCTTTTAC	GATGTTGGAC	3960
GATTTTTCT	TCTGATTGTT	GAATTCAAA	TCATGGCTC	ATTCCTT	CTTCTTGG	4020
ATATTCCTT	CAACACATT	CTTTATTTA	TTATACATTG	TGTCCTTTT	TTAGCTATTG	4080
CTGCTGTGT	TTTTTATTCT	ATTTACAGGA	TGATTTTAA	ACTGTC	AAAGTAGTGT	4140
TAACCTAAA	TAGGCTAAAT	GTGAACAAAT	AAAATACAGC	AAATACTCA	AAAAAAAAAA	4200
					AAAAAA	

BFG4 Protein sequence (SEQ_ID_NO:11)

Gene name: KIAA0882 protein; Unigene number: Hs.90419; Probeset Accession #: Z39762;
 Protein Accession #: BAA74905; Signal sequence: none; Predicted TM domains: 302-318; PFAM
 domains: TBC_domain: 135-347; Summary: a Type II membrane protein, likely localized to the
 peroxisome.

MTFLFANLKD	RDFLVQRISD	FLQQTTSKIY	SDKEFAGSYN	SSDDEVYSRP	SSLVSSSPQR	60
STSSDADGER	QPNLNGNSVP	TATQTLMTMY	RRRSPEEFNP	KLAKEFLKEQ	AWKIHFAEYG	120
QGICMYRTEK	TRELVLKGIP	ESMRGELWLL	LSGAINEKAT	HPGYYEDLVE	KSMGKYMLAT	180
EEIERDLHRS	LPEHPAFQNE	MGIAALRRLV	TAYAFRNPNI	GYCQAMNIVT	SVLLLYAKEE	240
EAFWLLVALC	ERMLPDPYNT	RUVGALVDQG	VPEELARDYV	PQLYDMQDL	GVISTISLSW	300
FLTLFLSVM	FESAVVVVDC	FFYEGIKVIF	QLALAVLDAN	VDKLLNCKDD	GEAMTVLGRY	360
LDSVTNKDST	LPPIPHLHSL	LSDDVEPYPE	VDFRLIRTS	YEKFGTIRAD	LIEQMRFKQR	420
LKVIQTLEDT	TKRNVVRTIV	TETSFTIDEL	EELYALFKAE	HLTSCYWGG	SNALDRHDPS	480
LPYLEQYRID	FEQFKGMFAL	LFPWACGTHS	DVLASRLFQL	LDENGDSLIN	FREFVSGLSA	540
ACHGDLTEKL	KLLYKMHVLP	EPSSDQDEPD	SAFEATQYFF	EDITPECTHV	VGLDSRSKQG	600
ADDGFVTVSL	KPDKGKRA	QNENRYLRLW	TPENKSKS	AKDLPKLNQG	QFIELCKTMY	660
NMFSEDNPNEQ	ELYTHAAATV	SLLLEIGEVG	KLFVAQPAKE	GGSGGSGPSC	HQGIPGVLF	720
KKGPQGPYVV	ESVEPLPASL	APDSEEHSLG	GQMEDIKLED	SSPRDNGACS	SMLISDDDTK	780
DDSSMSSY	LSAGSHEEDK	LHCEEIGEDT	VLVRSGQGTA	ALPRSTSLDR	DWAITFEQFL	840
ASLLTEPALV	KYFDKPVCM	ARITSAKNIR	MMGKPLTSAS	DYEISAMSG		

BCU7 DNA sequence (SEQ_ID_NO:12)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Nucleic Acid
 Accession #: n/a; Coding sequence: 1-573 (stop codon underlined)

TATTTTATTT	TCCAGGCTAA	AGCAAATGAA	AGTTTGCTGG	TATCAACACA	GCCTGCCATA	60
TTTTTCACAG	CATGCAACAA	TGGTGTAGG	ATAGCTATT	CTTACTGTAA	TTGCCAGAGG	120
CAGAAATGGT	CTGGGTATAA	GCTATTCAT	AAAAGCAGCT	TTAAATTGTC	AGTATTAAGG	180
TTTTCATGTG	GAAAGGTGTC	ATTCAAAAAA	AAAGTAATTG	GCATACATAT	TCCACATCAT	240
CGATCCCTCTC	TGTGGTGTAA	ATTTTTTAT	ATGACCACTG	GAAAAAATT	AATATTCTCA	300
CAATATAGT	TTGGGGCTT	CCATATCATC	AAAAGACTGA	AAAAATTATAA	TTTTAGAATT	360
AAACTGATGG	ATTCATTTAT	AGAATTATCT	GTGAGTTGTG	TAGACACAGT	CTTAATGTTT	420
CTGGTTATGA	CAGATAAGTT	TGCTAAAAA	ATGTGGATGA	AGCCATTATT	GTTATTATTG	480
TTATTGCTTC	TGTTCAGTTG	TCTAAGTATC	ATCCCCTCTG	TGGCCCATCA	CCGAGCAGAG	540
TTGCCCTACA	AATTCATTT	GGCAGCGCA	TAACATTATC	TTAAAAAGT	TATGAAACA	600
TTCATTGAA	AGTTCATGC	AGCTTCTAGCA	CAGAGTTGAC	CAAACACTGG	CGTAAGTTCA	660
ATTTCACAG	ATATTTGAA	TTGAACAAAT	AGAAATTTTT	CTCTAAATAT	ATACCTATGT	720
GAAACCAACT	TATCTGCATA	ATTTAACTCTA	ATACATATT	AAGCCAGTTT	AAGTGTGTTG	780
TGTTGATGCC	ATGCTTATCA	AATACATGCA	CAAGCTAAC	ATAATTGAA	TGGGTCTATG	840
AAGGAAAAAT	AATGCTTAA	CTTTGGTGT	GGTTCTTCTC	GTGTAGCCAT	ATACCCAGGC	900
TCTGCACTAT	CGAAGGATGC	AAATGTTGAC	ATAGATGGAA	GCTCTTACCT	ACCAAAGTGT	960
TTAGGAAAGGA	TAAGGTACA	TTTGCTTTAA	TTCTCAACAT	TATCTTGT	TTTATGTTT	1020
ATAAAAAATT	GTCTTATT	ATGCTGGTGA	AACGTATAAT	CACATCCAAT	TATTTGAACA	1080
CATGCAAAAT	AATTTTTAA	ATTATGTTAT	TGTTAAATT	TGACTTATGG	GAGATCAGTC	1140
AAAAACTTAG	AAGGTTAAC	ACTTCACTGA	TTATGTTGTC	TGAAACACG	TTACAAATTAC	1200
CACATATCCT	TGCTATAAGT	TTTGAAAGTT	CTTAGCAATT	AAAGTTTTT	TATTCAGTGT	1260
GAACGTGTCAG	TATCTTCT	GGTGTAAAT	GTATGGTGT	AAATGAATG	TTAGTGTG	1320
TCTGGTTAGT	ATATGCTCTT	TTATTCATTG	CTAAATT	TGTTATCCAT	TTGATTCTG	1380
ATTCAAGAAAT	ATCAATAAA	TCCTATGT	TTAATAC	TACCAAAAC	AGGCAAGTTA	1440
ACTCTGTTGT	TTTAATTCAA	CAGTCCAACA	TTATTTAGT	GTTACAGAGT	GTAAATATAT	1500
TTCTTGGGA	GT	TTTAAAT	TTTATAGC	TTGGAATGT	CCAAAGTC	1560
ATATCACCTA	AACTGGTTAG	ATTACTCTA	CAGCTAAAT	TATTGCAAGG	ACTGGCGCC	1620
TCTGGTGGTT	ATGAAGAAC	ATTCTTAATG	GCTACTTGA	CTACAGCAA	AGCCATTCT	1680
GTACCATAAA	AATTTGTTG	GCAATTATTAG	AAATTATCATA	TGTTTCTAC	ATCTGACAGC	1740
ACCTAAAATG	TTTGATAATA	TTAACATGTA	TCTAAGAGGA	AAAAAGAGT	AATATATTCT	1800
GGCACCCACT	TTCCTAGTAA	TGTTTCCAT	GATTTCCAG	TTCTGAGGCA	CTTATTAAG	1860
TGCTTTTTT	TTCTCTGAAAT	AATTAGGT	TGTTAAAATA	TATTTTAAAT	TTAGTTAGC	1920
TTTATAAAC	CAATTAGAAAT	TCACATTAA	TAACAGAGGT	ATAATTGT	CACTTCAGA	1980
AGTGATCATT	TATTTTATT	TAGCACAGGT	CATAAGAAA	ATATATAGAA	AAATAATCAA	2040

TTTCATATAT	AAAAGGATTA	TTTCTCCACC	TTTAATTATT	GGCCTATCAT	TTGTTAGTGT	2100
TATTTGGTCA	TATTATTGAA	CTAATGTATT	ATTCATTCA	AAGTCTTCT	AGATTTAAAA	2160
ATGTATGCAA	AAGCTTAGGA	TTATATCATG	TGTAACTATT	ATAGATAACA	TCCTAAACCT	2220
TCAGTTTAGA	TATATAATTG	ACTGGGTGTA	ATCTCTTTG	TAATCTGTTT	TGACAGATT	2280
CTTAAATTAT	GTTAGCATAA	TCAAGGAAGA	TTTACCTTGA	AGCACTTTCC	AAATTGATAC	2340
TTTCAAACCT	ATTTAAAGC	AGTAAACCT	TTCTATGAA	CTAAATCACA	TGCAAAACTC	2400
CAACCTGTAG	TATACATAAA	ATGGACTTAC	TTATTCCCTC	CACCTTCTCC	AGTGCCTAGG	2460
AATATTCTTC	TCTGAGCCCT	AGGATTGATT	CTATCACACA	GAGCAACATT	AATCTAAATG	2520
GTTTAGCTCC	CTCTTTTTC	TCTAAACAA	ATCAGCTAAT	AAAAAAAAAA	TTTGAGGGCC	2580
TAAATTATTT	CAATGGTTGT	TTGAAATATT	CAGTCAGTT	TGTACCTGTT	AGCAGTCTT	2640
CAGTTTGGGG	GAGAATTTAA	TACTGTGCTA	AGCTGGTGT	TGGATACATA	TTACAGCATC	2700
TTGTGTTTTA	TTTGACAAAC	AGAATTGTTG	TGCCATAATA	TTTGAGAAT	TAGAGAAGAT	2760
TGTGATGCAT	ATATATAAAC	ACTATTTTA	AAAAATATCT	AAATATGTCT	CACATATTAA	2820
TATAATCCTC	AAATATACTG	TACCATTTA	GATATTTTT	AAACAGATTA	ATTGGAGAA	2880
GTTTTATTCA	TTACCTAATT	CTGTGGAAA	AATGGTGCCT	CTGATGTTGT	GATATAGTAT	2940
TGTCACTGTC	TACATATATA	AAACCTGTGT	AAACCTCTGT	CCTATGAAAC	CATAACAAAT	3000
GTAGCTTTTT	AAAGTCCATT	GTATTGTTT	TTCTTCAAT	AAAAGAGTAT	ATTAATTGG	3060
TTGTTTTGAA						

BCU7 Protein sequence (SEQ ID NO:13)

Gene name: EST; Unigene number: Hs.98558; Probeset Accession #: AA428062; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 125-141, 154-170; PFAM domains: none; Summary: A type III membrane protein, highly overexpressed in breast cancer and prostate cancer; unknown function.

YFIFQAKANE	SLLVSTQPAI	FFTACNNGAR	IAISYCNCQR	QKWSGYKLFH	KSSFKLSQLR	60
FSCGKVFSKK	KVIGIHIPHH	RSSLWCFFY	MTRSKILIFS	QYRFWGFHII	KRLKNYNFRI	120
KLMDFIIELS	VSCVDTVLMF	LVMTDKFAQK	MWMKPLLLLL	LLLLFSCLSI	IPSVAHHAAE	180
LPYKFHLAAP						

BFA1 DNA sequence (SEQ ID NO:14)

Gene name: calsyntenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Nucleic Acid Accession #: NM_022131; Coding sequence: 11-2878 (start and stop codons underlined)

TGCTGCGAGG	<u>ATGCTGCC</u> GT	GGCGGGCTGTG	CTGGGTGCCG	CTCCTGCTGG	CGCTGGCGT	60
GGGGAGCGGC	AGCGCGGTG	GGGGGAGACAG	CCGGCAGCGC	CGCCTCTCG	CGGCTAAAGT	120
CATAAGCAC	AAAGCCATGGA	TCGAGACTTC	ATATCATGGA	GTCATAACTG	AGAACATGA	180
CACAGTCATT	TTGGACCCAC	CACTGGTAGC	CCTGGATAAA	GATGCACCGG	TTCCTTTGTC	240
AGGGGAAATC	TGTGCGTTCA	AGATCCATGG	CCAGGAGCTG	CCCTTGAGG	CTGTGGTGCT	300
CAACAAGACA	TCAGGAGAGG	GCCGGCTCCG	TGCCAAGAGC	CCCATTTGACT	GTGAGTTGCA	360
GAAGGAGTAC	ACATTCATCA	TCCAGGCC	TGACTGTGTT	GCTGGGGCCC	ACGAGACAGC	420
CTGGGAAAG	TCACACAAAG	CCGTGGTCCA	TATACAGGTT	AAAGATGTC	ACGAGTTTC	480
TCCCCACCTTC	AAAGAGCCAG	CCTACAAAGC	TGTGTGACG	GAGGGCAAGA	TCTATGACAG	540
CATTCTGCAG	GTGGAGGCCA	TTGACGAGGA	CTGCTCCCCA	CAGTACAGCC	AGATCTGCAA	600
CTATGAAATC	GTCACCACAG	ATGTGCTTT	TGCCATCGAC	AGAAATGGCA	ACATCAGGAA	660
CACTGAGAAG	CTGAGCTATG	ACAAACAAAC	CCAGTATGAG	ATCCTGGTGA	CCGCCTACGA	720
CTGTGGACAG	AAGCCCGCTG	CTCAGGACAC	CCTGGTCAG	GTGGATGTC	AGCCAGTTTG	780
CAAGCCTGGC	TGGCAAGACT	GGACCAAGAG	GATTGAGTAC	CAGCCTGGCT	CCGGGAGCAT	840
GCCCCCTGTT	CCCGACATCC	ACCTGGAGAC	GTGCGATGGA	GCCGTGTC	CCCTCCAGAT	900
CGTCACAGAG	CTGCAGACTA	ATTACATTGG	GAAGGGTTGT	GACCGGGAGA	CCTACTCTGA	960
GAATCCCTT	CAGAAGTTAT	GTGGAGCCTC	CTCTGGCATC	ATTGACCTCT	TGCCATCCCC	1020
TAGGCGCTGCC	ACCAACTGGA	CTGCGAGCT	GCTGGTGGAC	AGCAGTGGAGA	TGATCTTCAA	1080
GTTTGACGGC	AGGCGAGGTG	CCAAAATCCC	CGATGGGATT	GTGCGCAAGA	ACCTGACCGA	1140
TCAGTTTCACC	ATCACCAGT	GGATGAAACA	CGGCCCCAGC	CCTGGTGTGA	GAGCCGAGAA	1200
GGAAACCATC	CTCTGCACT	CAGACAAAC	CGAAATGAAC	CGGCATCACT	ATGCCCTGTA	1260
TGTGCAACAC	TGCCGCTCG	TCTTTCTCTT	GGCGAAGGAC	TTCGACCAAG	CTGACACCTT	1320
TCGCCCCCGC	GAGTTTCAACT	GGAAAGCTGA	TCAAGATTGT	GACAAAGAGT	GGCAACTACTA	1380
TGTCACTAAT	GTGGAGTTTC	CTGTGGTAAC	CTTATACATG	GATGGAGCAA	CATATGAACC	1440
ATACCTGGT	ACCAACGACT	GGGCCATTCA	TCCATCTC	ATAGCCATGC	AACTCACAGT	1500
CGGCCTTGT	TGGCAAGGAG	GAGAAGTCAC	CAAACACAG	TTTGTCTAGT	TCTTCATGG	1560
AAGCCTGGCC	AGTCTCACCA	TCCGCCCTGG	AAAATGGAA	AGCCAGAAGG	TGATCTCCTG	1620
CCTCGAGGCC	TGCAAGGAAG	GGCTGACAT	TAATTCCTTG	GAAAGCTTG	GGCAAGGAAT	1680
AAAGTATCAC	TTCACCCCT	CGCAGTCCAT	CCTGGTGTG	GAAGGTGACG	ACATTGGGAA	1740
CATTAACCGT	GCTCTCCAGA	AAAGTCTCTA	CATCAACTCC	AGGCAGTTC	CAACGGCGG	1800
TGTGCGGGCC	CTCAAAGTAT	CCTCCAAAGT	CCAGTGTGTT	GGGGAAGACG	TATGCATCAG	1860
TATCCCTGAG	GTAGATGCC	ATGTGATGGT	CCTCCAGGCC	ATCGAGCCC	GGATCACCC	1920
CCGGGGCACA	GACCACTCT	GGAGACCTGC	TGCCCAAGTT	GAAAGTGCC	GGGGAGTGAC	1980
CCTCTCCCT	GATATCAAGA	TTGTGAGCAC	CTTCGCCAAA	ACCGAACCCC	CCGGGGACGT	2040
GAAACACACA	GACCCCAAT	CAGAAGTCTT	AGAGGAAATG	CTTCATAACT	TAGATTCTG	2100
TGACATTTTG	GTGATCGAG	GGGACTTGG	CCCAAGGCAG	GAGTGTGTTG	AGCTCAACCA	2160
CAGTGAGCTC	CACCAACGAC	ACCTGGATGC	CACTAATTCT	ACTGCAGGCT	ACTCCATCTA	2220
CGGTGTGGGC	TCCATGAGCC	GCTATGAGCA	GGTGCTACAT	CACATCCGCT	ACCGCAACTG	2280
GCCTCCGGCT	TCCCTTGAGG	CCCGCGTTT	CCGGATTAAG	TGCTCAGAAC	TCAATGGGCG	2340
CTACACTAGC	AATGAGTCA	ACTTGGAGGT	CAGCATCCTT	CATGAAGACC	AAGTCTCAGA	2400
TAAGGAGCAT	GTCAATCATC	TGATTGTC	GCCTCCCTTC	CTCCAGTCTG	TCCATCATCC	2460
TGAGTCCCGG	AGTAGCATCC	AGCACAGTTC	AGTGGTCCCA	AGCATTGCA	CAGTGGTCAT	2520

CATCATCTCC	GTGTGCATGC	TTGTGTTGT	CGTGGCCATG	GGTGTGTACC	GGGTCCGGAT	2580
CGCCCCACCA	CACTTCATCC	AGGAGACTGA	GGCTGCCAAG	GAATCTGAGA	TGGACTGGGA	2640
CGATTCTCGC	CTGACTATCA	CACTCAACCC	CATGGAAA	CATGAAGGAC	CAGGGCATGG	2700
GGAAAGATGAG	ACTGAGGGAG	AAGAGGAGGA	AGAAGCCAG	GAAGAAATGA	GCTCCAGCAG	2760
TGGCTTGAC	GACAGCGAAG	AGGAGGAGGA	GGAGGAAGGG	ATGGGCAGAG	GCAGACATGG	2820
GCAGAAATGGA	GCCAGGCAAG	CCCAGCTGA	GTGGGATGAC	TCCACCCCTCC	CCTACTAGTG	2880
CCCAGGGTC	TGCTGCCCTGG	CCCACATGTC	CTTTTGTA	ACCCCTGACCC	AGTGTATGCC	2940
CATGTCTATC	ATACCTCAC	TCTGATGTCT	GTGACATGTC	GGGGAGGCC	TTCTCCAGCT	3000
TCTCTGGAGCC	CACCCCTTAA	GCCTTGGGCA	CTCCCTGTG	TTCATCCATG	GGGAAGTTCC	3060
AAGAAGCCCA	GCATGGCCAT	CACTGAGGAC	TTCAGGGTAG	ACTTTGTCCT	GTAGCCTCCA	3120
CTTCTGCCCT	AAGTTCCCCA	GCATCTGAC	TACCTGTC	CAGAGTTG	CTTTGTTTT	3180
TCCTGCAGGG	AAGAAGGCC	ACCTTGTTG	CACTCACCTC	CCCAGGCTCA	GAGTCCCCAA	3240
GGCCCTGGGG	TTCCAACCTCA	CTGTGGTCT	CCTCCACACA	GACCAGTAGG	TTCTCCCTATG	3300
CTGACTCCAG	GTTGCTTCAT	ACAAGGAGGG	TGGTTGAACT	TCACACACAG	AAGGTCTTAG	3360
TGCTTAACAG	TTTAAAGGAA	AGTCCTGTG	GAGGAGGAC	TAAGTTTACA	GGGAAAGGTA	3420
CACACATTCT	CTCTCTCTG	CTATCTGTG	CATCTAGTT	CCCCAGCTTG	GAGAGCCTT	3480
CCCTTGTGTT	CTTCTGAGG	CCATATAAGC	TTATAAGAAA	AGTCCCAAAC	CAAGAATAGG	3540
TCCTTGGCCA	CAAGCAGGGT	CTGATCCCC	ATCAGAGCTA	TCTGAGCCTG	CCTGCTGGG	3600
CACCTGCTGC	AACCATCGAG	CTACCTGCC	AGGGGCACTC	ACCAACACAG	ACCACAGGGC	3660
CCAGGAGGCA	TTCCACACAG	GCACGCC	AGGACAACAC	ACAAGGACA	GTCACAACAA	3720
GGACAACAG	GACACAAAC	AACACACAC	AAGGACAGTC	ACAACAGCC	TAGAGCCAGA	3780
AAGCAGATGG	AAATGCTAAT	GAGGCTAAC	GTAGGCTCA	TGGTGGGTG	AGTGGGGGTG	3840
GCTGGGCTCC	CCCAGGACAG	AGGGGACCT	GAGGTTGGCA	AGGCTCTC	CACTCAGCCT	3900
TATGGTCCCT	TATCTCCTAT	CTTCCCTCTT	GAGAAAATAC	ACGCTTCTG	CATGTATTAG	3960
AAACGCACGA	GCTCCACCAA	GTCTCAATG	AAAGTTGAA	ATTTAACTGC	AAGGAATTAG	4020
AAGCATATT	GCAATCATG	CAGCTCTTC	TTCTTCTG	TCATAAAAAGG	AGGAACACTT	4080
TAGATAGAGG	SCAATATAT	CTGAAAACCT	AAATTCTTTC	TTTTTTGAT	AAGGAATCT	4140
TTTCCATCTC	CATCCTAAC	TGCACAAACCT	GTGAAGAGAA	TTGTTTCTAT	AGTAACGTG	4200
CTGTGATCTT	TTGTGGCCAA	GAGAATAGCA	GGCAAGAATT	AGGGCCTTG	CAGAATTTC	4260
ACGAAGCTCT	GAGAACATGT	TTGTTTCAA	TGCTCTGATTC	CTCTTGTCA	TCAATGTGA	4320
TGCTCTGTC	CCATCCTTCA	CTCCTCTCA	AGTCACACC	AATTGTTTG	GCACAGGCAC	4380
AGAGCTGGTC	CCTAGTTAAG	TGGCATTAT	GTTAAAAAAA	A		

BFA1 Protein sequence (SEQ ID NO:15)

Gene name: calsyntenin-2; Unigene number: Hs.7413; Probeset Accession #: R46025; Protein Accession #: NP_071414; Predicted Signal sequence: 1-20; Predicted TM domains: 832-848; PFAM domains: cadherin domains: 48-151, 165-254; Summary: A type I membrane protein; a member of the calsyntenin family; is related to the FAT tumor suppressor; is likely an adhesion molecule important in mammalian developmental processes and cell communication.

MLPGRLCWVP	LLLALGVGSG	SGGGGDSRQR	RLLAALKVNH	KPWIETSYHG	VITENNDTVI	60
LDPPLVALDK	DAPVPFAGEI	CAFKIHQQL	PFEAVVLNKT	SGEGRRLRAKS	PIDCELOKEY	120
TFI IQAYDCG	AGPHETAWKK	SHKAVVHIVQ	KDVNEFAPTF	KEPAYKAVVT	EGKIYDSILQ	180
V EAIDEDCSP	QYSQICNYEI	VTTDVPFAID	RNGNRNTEK	LSYDKQHQYE	ILVTAYDCGQ	240
KPAAQDQTLVQ	VDVKPVCKPG	WQDWTKRIEY	QPQSGSMPLF	PSIHLLETCDG	AVSSLQIVTE	300
LQTNYIGKGC	DRETYSEKSL	QKLCGASSGI	IDLILPSPSAA	TNWTAQALLV	SSEMIFKFDFG	360
RQGAKIPDG	VPKNLTQFT	ITMWMKHGPS	PGVRAEKETI	LCNSDKE	RHYALYVHN	420
CRLVFLLRKD	FDQADTFRPA	EFHWKLQDQC	DKEWHYVVIN	VEFPVVTLYM	DGATYEPYLV	480
TMDWPIHPSH	IAQLTVGAC	WQGGEVTKPQ	FAQFFHGSLA	SLTIRPGKME	SQKVISCLQA	540
CKEGLDINSL	ESLGQGIKYH	FNPSQSILVM	EGDDIGNIR	ALQKVSYAINS	RQFPTAGVRR	600
LKVSSKVQCF	GEDVCISIPE	VDAYVMVLAQ	I EPRITLRTG	DHFWRPAAQF	ESARGVTLFP	660
DIKIVSTFAK	TEAPGDVKT	DPKSEVLEEM	LHNLDPCDIL	VIGGDLDPHQ	ECLELNHSEL	720
HQRHLDATNS	TAGYSIYVG	SMSRYEQVLH	HIRYRNWRPA	SLEARRFRIK	CSELNGRYTS	780
NEFNLEVSIL	HEDQVSDEK	VNHLLIVQPPF	LQSVHHPESR	SSIQHSSVVP	SIATVIIIS	840
VCMLVFFFAM	GVYRVRIAHQ	HFIQETEAAC	ESEMDWDSSA	LTITVNPMEK	HEPGHGDE	900
TEGEEEEEAE	EEMSSSSGSD	DSEEEEEE	MGRGRHGQNG	ARQAQLEWDD	STLPY	

BFG7 DNA sequence (SEQ ID NO:16)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Nucleic Acid Accession #: n/a; Coding sequence: <1-906 (stop codon underlined)

CGGGTCGACC	CACCGCTCCG	GGGAGAAAGG	ATGGCCGCC	TGGCGGCCG	GTTGGCTCTG	60
CTAGCTGGGG	CAGCGGCCT	GGCGAGCGGC	TCCCAGGGCG	ACCGTGAGCC	GTTGTACCGC	120
GACTGCTAC	TGCAGTGC	AGAGCAGAAC	TGCTCTGGG	GCGCTCTGAA	TCACTTCCGC	180
TCCCGCCAGC	CAATCTACAT	GAGTCTAGCA	GGCTGGACCT	GTCGGGACGA	CTGTAAGTAT	240
GAGTGTATGT	GGGTCAACCGT	TGGCTCTAC	CTCCAGGAAG	GTCACAAAGT	GCCTCAGTTC	300
CATGGCAAGT	GGCCCTTCTC	CCGGTCTCTG	TTCTTCAAG	AGCCGGCATC	GGCCGTGGCC	360
TCGTTTCTCA	ATGGCCTGGC	CAGCCTGGT	ATGCTCTG	GCTACCGCAC	CTTCGTGCCA	420
GCCTCCCTCCC	CCATGTACCA	CACCTG	GCCTTCG	GGGTGTCCT	CAATGCATGG	480
TTCTGGTCCA	CAGTYTTCCA	CACCA	GAGCTG	CTTG	GGACTACTTC	540
TGTGCTCCA	CTGTCACTCT	ACACTCA	TACCTG	GCCTCAGCCT	CATCCCTTC	600
GACTATGGCT	ACAACCTGGT	GGCAAC	GCTATT	TGGTCAACGT	GGTGTGGTGG	660
CTGGCCTGGT	GCCTGTGGAA	CCAGCGGCCG	CTGCTC	TGCGCAAGTG	CGTGGTGGT	720
GTCTTGCTGC	TGCA	GGCTG	ACTTCCCACC	GCTCTTCTG	780	
GTCCTGGATG	CCCCATGCAT	CTGGCACATC	AGCACCATCC	CTGTCCACGT	CCTCTTTTTC	840
AGCTTTCTGG	AAGATGACAG	CCTGTACCTG	CTGAAGGAAT	CAGAGGACAA	GTTCAAGCTG	900

GAATGAAAGAC	CTTGGAGCGA	GTCTGCCCA	GTGGGGATCC	TGCCCCGCC	CTGCTGGCCT	960
CCCTTCTCCC	CTCAACCCCT	GAGATGATTT	TCTCTTTC	ACTTCTTGAA	CTTGGACATG	1020
AAGGATGTGG	GCCCCAGAAC	ATGTGCCAG	CCCACCCCT	GTGGCCCTC	ACCAGCCTTG	1080
GAGTCTGTC	TAGGAAAGGC	CTCCCAGCAT	CTGGGACTCG	AGAGTGGGCA	GCCCCCTCTAC	1140
CTCCTGGAGC	TGAACCTGGGG	TGGAACATGAG	TGTCTCTTA	GCTCTACCGG	GAGGACAGCT	1200
GCCCTTTCC	TCCCCATCAG	CCTCCCTCCC	ACATCCCCAG	CTGCCCTGGC	GGGTCTGAA	1260
GCCCTCTGTC	TACCTGGGAG	ACCAGGGACC	ACAGGCCCTA	GGGATACAGG	GGGTCCCCCTT	1320
CTGTTACAC	CCCCCACCC	CCTCCAGGAC	ACCACTAGGT	GGTGCTGGAT	GCTTGTCTT	1380
TGGCCAGGCCA	AGGTTCACCG	CGATTCTCCC	CATGGGATCT	TGAGGGACCA	AGCTGCTGGG	1440
ATTGGGAAGG	AGTTTCACCC	TGACCRRTGC	CCTAGCAGG	TTCCCGAGGAG	GCCTCACCAT	1500
ACTCCCTTC	AGGGCCAGGG	CTCCAGCAAG	CCCCAGGGCAA	GGATCTGTG	CTGCTGTCTG	1560
GTTGAGAGCC	TGCCACCGTG	TGTCGGGAGT	GTGGGCCAGG	CTGAGTGCA	AGGTGACAGG	1620
GCCGTGAGCA	TGGGCCTGGG	TGTGTGTGAG	CTCAGGCACT	AGGTGCGCAG	TGTGGAGACG	1680
GGTGTTGTCG	GGGAAGAGGT	GTGGCTTCA	AATGTTGTGT	GTGCAAGGGG	TKGGTGTGTT	1740
AAGCCTGGGT	TAGGGAAACG	TGTGTGCGCG	TGCTGGTGGG	CATGTGAGAT	GAGTGACTGC	1800
CGGTGAATGT	GTCCACAGGT	GAGAGGTTGG	AGCAGGATGA	GGGAATCTC	TCACCATCAA	1860
TAATCACTTG	TGGAGGCCCA	CTTGGCCCAA	GACGCCACCT	GGGCGGACAG	CAGGAGCTCT	1920
CCATGGCCAG	GCTGCCTGTG	TGCATGTTCC	CTGCTGGTG	CCCCCTTGCC	CGCCTCTGC	1980
AAACCTCACA	GGGTCCCCAC	ACAAACAGTC	CCTCCAGAAG	CAGCCCCCTC	GAGGCAGAGG	2040
AAGGAAATG	GGGATGGCTG	GGGCTCTCTC	CATCCTCTT	TTCTCCTTGC	CTTCGCATGG	2100
CTGGCCTTC	CCTCCAAAAC	CTCCATTCCC	CTGCTGCCAG	CCCCCTTGCC	ATAGCCTGAT	2160
TTTGGGGAGG	AGGAAGGGGC	GATTTGAGGG	AGAAGGGGAG	AAAGCTTATG	GCTGGGTCTG	2220
GTTCCTTCCC	TTCCCAGAGG	GTCTTACTGT	TCCAGGGTGG	CCCCAGGGCA	GGCAGGGGCC	2280
ACACTATGCC	TGCGCCCTGG	AAAGGTGAC	CCCTGCCATT	TACCAAGCAGC	CCTGGCATGT	2340
TCCTGCCCCA	CAGGAATAGA	ATGGAGGGAG	CTCCAGAAC	TTTCCATCCC	AAAGGCAGTC	2400
TCCGTGGTTG	AAGCAGACTG	GATTTTGCT	CTGCCCCCTGA	CCCCCTTGCC	CTCTTTGAGG	2460
GAGGGAGGCT	ATGCTAGGAC	TCCAACCTCA	GGGACTCGGG	TGGCCTGCG	TAGCTTCTT	2520
TGATACTGAA	AACTTTAAG	GTGGGAGGGT	GGCAAGGGAT	GTGCTTAATA	AATCAATTCC	2580
AAAGCCTCAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAA			

BFG7 Protein sequence (SEQ ID NO:17)

Gene name: EST; Unigene number: Hs.91668; Probeset Accession #: Z40805; Protein Accession #: n/a; Signal sequence: none; Predicted TM domains: 117-133, 179-195, 211-227, 235-251, 266-282, 296-312; PFAM domains: none; Summary: A type III membrane protein of unknown function; is adjacent to HER2 on the genome, and its overexpression in breast cancer is highly correlated with HER2 expression; may be used to predict HER2 overexpression and amplification.

RVDPRVRGER	MAGLAARLVL	LAGAAALASG	SQGDREPVYR	DCVLQCEEQN	CSGGALNHFR	60
SRQPIYMSLA	GWTCRDDCKY	ECMWVTVGLY	LQEGHKVPQF	HGKWPFSRFL	FFQEPAASAVA	120
SFLNGLASLV	MLCRYRTFVP	ASSPMYHTCV	AFAWVSLSNAW	FWSTVFHTRD	TDLTEKMDYF	180
CASTVILHSI	YLCCVVRTVGL	QHPAVVSAAFR	ALLLLMLTVH	VSYLSSLIRFD	YGYNLVANVA	240
IGLVNVWWL	AWCLWNQRRL	PHVRKCVVVV	LLLQGLSLL	LLDFPPLFWV	LDAHAIWHIS	300
TIPVHVLFFS	FLEDDSLYLL	KESEDKFKLD				

BCN4 DNA sequence (SEQ ID NO:18)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Nucleic Acid Accession #: n/a; Coding sequence: 143-874 (start and stop codons underlined)

GGGAGGGAGA	GAGGCGCGCG	GGTGAAGAGC	GCATTGATGC	AGCCTGCC	GGCCTCGGAG	60
CGCGGCGGAG	CCAGACGCTG	ACCACGTTCC	TCTCCTCGGT	CTCCTCCGCC	TCCAGCTCCG	120
CGCTGCCCGG	CAGCCGGAG	CCATCGCGAC	CCAGGGCCCC	GCCGCTCTCC	CGCAGCGGCT	180
CCCGGGCCCT	CTGCTGCTCC	TGCTGCTGCA	GCTGCCCGCG	CCGTCGAGCG	CCTCTGAGAT	240
CCCCAAGGGG	AAGCAAAGG	CGCACGCTCCG	GCAGAGGGAG	GTGGTGGACC	TGTATAATGG	300
AATGTGCTTA	CAAGGGCAG	CAGGAGTGCC	TGGTCGAGAC	GGGAGCCCTG	GGGCAATGG	360
CATTCCGGGT	ACACCTGGGA	TCCCAGTCG	GGATGGATT	AAAGGAGAAA	AGGGGAATG	420
TCTGAGGGAA	AGCTTTGAGG	AGTCTCGGAC	ACCCAAC	AAGCAGTGT	CATGGAGTTC	480
ATTGAATTT	GGCATGATC	TTGGGAGTAT	AGCTGGAGTGT	ACATTACAA	AGATGCGTT	540
AAATAGTGCT	CTAACAGAGTT	TGTTCACTGG	CTCACTTCGG	CTAAATGCA	GAAATGCATG	600
CTGTCAGCGT	TGGTATTTC	CATTCAATGG	AGCTGAATGT	TCAGGACCTC	TTCCCATGTA	660
AGCTATAATT	TATTGGACC	AAGGAAGCCC	TGAAATGAAT	TCAACAATT	ATATTGATCG	720
CACTTCTTCT	GTGGAAGGAC	TTTGTAAGG	AATTGGTGCT	GGATTAGTGG	ATGTTGCTAT	780
CTGGGTGGC	ACTTGTTCAG	ATTACCCAAA	AGGAGATGCT	TCTACTGGAT	GGAAATTCAGT	840
TTCTCGATC	ATTATGAAAG	AACTACCCAA	<u>AT</u> AAATGCTT	TAATTTCAT	TTGCTACCTC	900
TTTTTTATT	ATGCCTTGG	ATGGTCACT	TAATGACAT	TTAAATTAAG	TTTATGTATA	960
CATCTGAATG	AAAAGCAAAG	CTAAATATGT	TTACAGACCA	AAGTGTGATT	TCACACTGTT	1020
TTAAATCTA	GCATTATCA	TTTGCTTCA	ATC AAAAGTG	TTTTCAATAT	TTTTTTAGT	1080
TGGTTAGAAT	ACTTTCTCA	TAGTCACATT	CTCTCAACCT	ATAAATTGGA	ATATTGTTGT	1140
GGTCTTTGT	TTTTCTCTT	AGTATAGCAT	TTTAAAAAAA	ATATAAAAAGC	TACCAATCTT	1200
TGTACAATT	GTAAATGTTA	GAATTTTTT	TTATATCTGT	AAATAAAAAA	TTATTCCAA	1260
CAACCTTAAA	AAAAAAAAAA	AAAAAAA				

BCN4 Protein sequence (SEQ ID NO:19)

Gene name: ESTs; Unigene number: Hs.283713; Probeset Accession #: F13673; Protein Accession #: n/a; Predicted Signal sequence: 1-30; TM domains: none; PFAM domains: none; Summary: a secreted protein; has a mouse orthologue (see sequence below).

MRPQGPAAASP QRLRGLLLLL LLQLPAPSSA SEIPKGKQKA QLRQREVVDL YNGMCLQGPA 60
GVPGRDGSPG ANGIPGTPGI PGRDGFKGEK GECLRESFEE SWTPNYKQCS WSSLNYGIDL
GKIAECTFTK MRSNSALRVL FSGSLRLKCR NACCQRWYFT FNGAECSGPL PIEAIYLDQ
GSPEMNSTIN IHRRTSSVEGL CEGIGAGLVD VAIWVGTCS DYPKGDASTGW NSVSRIIIIEE
LPK

Mouse BCN4 Protein sequence (SEQ ID NO:20)
Gene name: ESTs; Unigene number: Mm.41556

XXXXAAPPQL LLGLFLVLLL LLQLSAPSSA SENPKVKQKA LIRQREVVDL YNGMCLQGPA 60
GVPGRDGSPG ANGIPGTPGI PCQDGFKGEK GECLRESFEE SWTPNYKQCS WSSLNYGIDL
GKIAECTFTK MRSNSALRVL FSGSLRLKCR NACCQRWYFT FNGAECSGPP PIEAIXXXX
XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXSD YPKGDAYTGW DSVSRIIIIEE
LPK